Volume 7 DECEMBER, 1936 Number 12

PHYSICS

A Journal of General and Applied Physics

WITH the next issue, *Physics* will appear as the *Journal of Applied Physics*, a publication of the American Institute of Physics.

For the past year a committee of the American Physical Society, under the leadership of Dr. M. J. Kelly has been making a thorough study of the ways in which the Society can be of service to those physicists who not only are interested in physics as a fundamental science but also are actively engaged in applying that scence to other sciences and to industry.

One of the most immediate matters to which the committee turned its attention was the publication needs of these applied physicists, and the manner in which the journal *Physics* should be developed to meet their needs. In this connection the committee sought and obtained the opinions and points of view of a large number of representative physicists in industry. In addition it had the guidance of the Advisory Council on Applied Physics of the American Institute of Physics. After careful study the committee evolved and the Council of the American Physical Society approved a plan of publication which you will see developed during the coming year in the *Journal of Applied Physics.*

Under the capable editorship of Dr. Elmer Hutchisson and a board of associate editors drawn from the various specialties in applied physics the *Journal of Applied Physics* will become the professional journal for the applied physicist. In it will appear not only research papers in applied physics as heretofore but also editorials, news items, review articles, book reviews, and advertisements of interest to applied physicists.

The *Journal of Applied Physics* in its new and attractive format will appeal to all physicists, engineers, and chemists who are applying physics to their special problems. The character of the journal is just being formed and I know that I speak for the new management when I urge that you take a proprietary interest in its development.

JOHN T. TATE
Editor

The Distribution of Relaxation Times in Typical Dielectrics*

WILLIAM A. YAGER, *Bell Telephone Laboratories, New York City*

(Received September 8, 1936)

K. W. Wagner's treatment of the distribution of relaxation times in dielectrics is reviewed; the effect of the density of distribution upon the frequency variation of the dielectric constant and dielectric loss factor is discussed; and a graphical method of evaluating the constants of Wagner's equation from experimental data is described. These constants have been evaluated for a number of typical dielectrics, and the frequency variations of the dielectric constant and dielectric loss factor as computed from Wagner's equations and from the simple equations for a single relaxation time are compared with the observed behavior. The quantitative correlation obtained between Wagner's equations and the experimental data investigated is viewed as a confirmation of Wagner's theory of a statistical distribution of relaxation times in dielectrics.

INTRODUCTION

NUMEROUS investigators have found that the well-known simple absorption equations for the time variation of the reversible absorption current and for the frequency variation of the dielectric constant ϵ', and the dielectric loss factor ϵ'', are not in accord with their experimental data. It is generally agreed that the failure of these simple absorption equations to explain the experimental data is to be ascribed to the polydisperse nature of the systems investigated. The form of the absorption equations is governed by the form of the relaxation function in the fundamental equations of dielectric absorption. The simple absorption equations correspond to a relaxation function consisting of a single negative exponential term. Although the basic theories of dielectric absorption all lead to a relaxation function of this form, they regard a dielectric as being a monodisperse system, that is, a system in which the factors giving rise to the absorption are nonvariant within the medium. Such a system is characterized by a single relaxation time. Many dielectrics, however, are polydisperse systems, that is, the factors giving rise to the absorption vary within the medium so that the dielectric is characterized not by a single relaxation time but by what may be termed a distribution of relaxation times.

The concept of a distribution of relaxation times was first introduced by von Schweidler[1] in 1907. He represented the relaxation function as a summation of a large number of negative exponential terms, there being one term for each different relaxation time. On the assumption that the relaxation times vary continuously from 0 to ∞, he then expressed the relaxation function as a definite integral involving a distribution function. Instead of developing a distribution function from theoretical considerations, he worked backwards and determined mathematically the distribution function that would give the empirical relaxation function,

$$\Phi(t) = Bt^{-n},$$

first proposed by J. Hopkinson.[2]

In 1913, K. W. Wagner[3] proposed that the distribution of relaxation times in dielectrics is governed by a probability function and derived the corresponding absorption equations for the time variation of the reversible absorption current and the frequency variation of ϵ' and ϵ''. In 1914, he obtained quantitative agreement between the theoretical time variation of the reversible absorption current and that actually observed in various samples of balata.[4] However, the validity of the equations for the frequency variation of ϵ' and ϵ'' apparently has never been investigated. In 1926, Karapetoff[5] proposed a distribution function of the form $F(\alpha) = H\alpha^{m-1}e^{-n\alpha}$ in which H, m and n are constants independent of α, and α is the reciprocal of the relaxation time. However, the integrals resulting from the introduction of this distribution function into the equations for the frequency variation of ϵ' and ϵ''

* Presented at the eighth annual meeting of the Committee on Electrical Insulation of the National Research Council at Pittsfield, Mass., Oct. 17, 1935.

[1] E. von Schweidler, Ann. d. Physik **24**, 711 (1907).

[2] J. Hopkinson, Phil. Trans. **166**, 489 (1876); Phil. Mag. **2**, 314 (1876); Phil. Trans. **167**, 599 (1877); Proc. Roy. Soc. **25**, 496 (1876).

[3] K. W. Wagner, Ann. d. Physik **40**, 817 (1913).

[4] K. W. Wagner, Archiv. f. Elektrotechnik **3**, 83 (1914).

[5] V. Karapetoff, J. A.I.E.E. **45**, 236 (1926).

lead to infinite series when m is not an integer and hence very little simplification, if any, results. The applicability of this distribution function has not been investigated.

In 1932, Whitehead and Baños[6] obtained excellent agreement between a charge and discharge—time relationship determined experimentally on a specimen of dry paper at 60°C and 1500 volts and values computed from a relaxation function consisting of three negative exponentials in which the six constants were evaluated by the method of successive residuals as outlined by J. Lipka.[7] Although this procedure suffices for the predetermination of the $a-c$ characteristics of dielectrics from suitable charge and discharge curves, as applied by Whitehead and Baños, it cannot be viewed as a confirmation of von Schweidler's theory but rather as a means of obtaining an empirical equation by the choice of a sufficient number of constants. Wagner's equations on the other hand involve but four constants all of which have physical significance. The only constant not appearing in the simple absorption equations is b which governs the density of distribution of the relaxation times. Hence, quantitative agreement between Wagner's equations and experimental data may be viewed not only as a confirmation of von Schweidler's theory concerning a distribution of relaxation times but also as a confirmation of Wagner's proposal that this distribution is governed by a probability function.

Recently several experimenters[8–14] investigating the mechanism of dielectric absorption in various dielectrics have found that the simple equations for the frequency variation of ϵ' and ϵ'' are inadequate to explain their experimental results which indicate broader regions of absorption and lower dielectric loss factor maxima than can be accounted for by the simple absorption equations. It is generally agreed that these discrepancies between experiment and theory are to be attributed to the polydisperse nature of the systems investigated. In this connection, Race[8] has presented an illustration in which he considers a hypothetical liquid dielectric composed of equal numbers of polar molecules of five different sizes. The change in the measured ϵ' would represent the sum of the contributions of molecules of all sizes. The loss factors, however, would have their maxima at different frequencies since the relaxation time depends upon the particle size. Therefore, the sum of the loss curves would have a maximum ordinate (ϵ_m'') considerably less than would have been obtained if the molecules had been all of one size. For the assumed case, the maximum ordinate of the sum of the five loss curves is only about 62 percent of the maximum ordinate for the curve for which all the particles are assumed to be of one size only. The resulting ratio is nearly the same as that which he found between the observed and calculated maximum values of ϵ''.

Although these authors[8–14] have pointed out that their data indicate polydisperse systems, no mention has been made as to the probable distribution of relaxation times therein nor has the validity of Wagner's equations for the frequency variation of ϵ' and ϵ'' for a statistical distribution of relaxation times been investigated. The object of this paper is, therefore, to present a method of examining experimental data on anomalous dispersion and dielectric loss on the basis of Wagner's hypothesis and to point out the quantitative agreement obtained between the experimental and theoretical curves for the frequency variation of ϵ' and ϵ'' for various typical dielectrics.

Theoretical Development

In order to comprehend fully the significance of Wagner's equations, it is desirable to review briefly the development of the fundamental equations of dielectric absorption. The time variation of the reversible absorption current, i_a, under continuous potential is characterized by the expression,

$$i_a = C_\infty E \Phi(t), \tag{1}$$

[6] J. B. Whitehead and A. Baños, Trans. A.I.E.E. **51**, 392 (1932).
[7] J. Lipka, *Graphical and Mechanical Computation* (John Wiley & Sons, 1918).
[8] H. H. Race, J. Phys. Chem. **36**, 1928 (1932).
[9] D. W. Kitchin, J. Ind. and Eng. Chem. **24**, 549 (1932).
[10] A. H. White and S. O. Morgan, (a) Physics **2**, 313 (1932); (b) J. Frank. Inst. **216**, 635 (1933).
[11] J. L. Oncley and J. W. Williams, Phys. Rev. **43**, 341 (1933).
[12] H. Rieche, Zeits. f. Physik **95**, 158 (1935).
[13] G. M. L. Sommerman, Rev. Sci. Inst. **5**, 341 (1934); J. Frank. Inst. **219**, 433 (1935).
[14] A. L. Ferguson, L. O. Case and G. H. Evans, J. Chem. Phys. **3**, 285 (1935).

where C_∞ denotes the geometric capacitance; E, the applied voltage, and $\Phi(t)$, the relaxation function. Eq. (1) embodies Curie's laws which state that the reversible absorption current is directly proportional to the applied voltage and to the dimensions of the condenser. If these laws are obeyed, Hopkinson's principle of superposition may be applied for the determination of the reversible absorption current under any continuous variation of the applied voltage $E(t)$. According to von Schweidler, i_a, measured at any time, t_m, then becomes

$$i_a = C_\infty \int_{-\infty}^{t_m} \frac{dE(t)}{dt} \Phi(t_m - t)dt, \tag{2}$$

where t denotes the time at which a given change in voltage occurred. $t_m - t$ then gives the elapsed time between the given change in voltage and the measurement of i_a. Denoting this elapsed time by u,

$$i_a = -C_\infty \int_0^\infty \frac{dE(t_m - u)}{du} \Phi(u)du. \tag{3}$$

For a sinusoidal voltage,

$$\frac{dE(t_m - u)}{du} = -E_0\omega \cos \omega(t_m - u),$$

where $\omega = 2\pi \times$frequency. Eq. (3) then becomes

$$i_a = C_\infty \omega E_0 \int_0^\infty \cos \omega(t_m - u)\Phi(u)du.$$

But $\cos \omega(t_m - u) = \cos \omega t_m \cos \omega u + \sin \omega t_m \sin \omega u$. Hence,

$$i_a = C_\infty \omega E_0 \Big[\cos \omega t_m \int_0^\infty \Phi(u) \cos \omega u \, du + \sin \omega t_m \int_0^\infty \Phi(u) \sin \omega u \, du \Big]. \tag{4}$$

From this equation, the dielectric loss factor is found to be

$$\epsilon'' = \epsilon_\infty \int_0^\infty \Phi(u) \sin \omega u \, du \tag{5}$$

and the increase in dielectric constant due to dielectric absorption is found to be

$$\epsilon' - \epsilon_\infty = \epsilon_\infty \int_0^\infty \Phi(u) \cos \omega u \, du. \tag{6}$$

Eqs. (1), (5) and (6) may be considered the fundamental equations of dielectric absorption since they are perfectly general and apply regardless of the mechanism of dielectric absorption or form of the relaxation function provided that Curie's laws are obeyed. Thus, if the relaxation function, $\Phi(u)$, is known, the time variation of the reversible absorption current and the frequency variation of the dielectric constant and dielectric loss factor may be determined.

In the theories of Pellat[15] and of Maxwell[16] both as originally proposed for a two-layer dielectric and as later extended by Wagner[17] for a disperse system, the relaxation function is a negative exponential of the form

$$\Phi(t) = (k/\tau)e^{-t/\tau}, \tag{7}$$

where k is the absorption constant and τ is the relaxation time. Substituting (7) in Eqs. (1), (5) and (6),

$$i_a = \frac{C_\infty k}{\tau} E e^{-t/\tau},$$

$$\epsilon'' = \frac{\epsilon_\infty k}{\tau} \int_0^\infty e^{-t/\tau} \sin \omega t \, dt, \tag{1a}$$

$$\epsilon' = \epsilon_\infty \Big[1 + \frac{k}{\tau} \int_0^\infty e^{-t/\tau} \cos \omega t \, dt \Big].$$

From formulae 506 and 507 of Pierce's Table of Integrals,

$$\epsilon'' = (\epsilon_\infty k\omega\tau/1 + \omega^2\tau^2), \tag{5a}$$

and

$$\epsilon' = \epsilon_\infty[1 + k/(1 + \omega^2\tau^2)]. \tag{6a}$$

Any mechanism of dielectric absorption in which the relaxation function is a simple negative exponential leads to equations of the form of (1a), (5a) and (6a). Debye's theory[18] of polar molecules likewise leads to equations of the form of (5a) and (6a) although no specific relaxation function is given. However, it seems reasonable

[15] H. Pellat, Ann. Chim. Phys. **18**, 150 (1899).
[16] J. C. Maxwell, *Electricity and Magnetism*, Vol. 1 (Oxford, 1892), p. 328.
[17] K. W. Wagner, Archiv. f. Elektrotechnik **2**, 371 (1914).
[18] P. Debye, *Polar Molecules* (Chemical Catalog Company, 1929).

to conclude that inasmuch as this theory leads to equations of the same form as (5a) and (6a), the relaxation function, if it were explicitly defined, would be a simple negative exponential.

Although the various mechanisms of dielectric absorption all lead to a simple exponential relaxation function, such a simple function is not in accord with experimental data. In order to account for the disagreement between experiment and theory, von Schweidler introduced the concept of a distribution of relaxation times by extending Pellat's theory. He regarded the charged particles in a dielectric as being of two types, the motion of one being oscillatory and the motion of the other being highly damped and hence aperiodic. Upon the sudden application of an external field, the displacement of the particles of the oscillatory type is instantaneous and constitutes the normal displacement whereas the displacement of the aperiodic particles is a function of time,

$$D_v(t) = D_0 k(1 - e^{-t/\tau})$$

and constitutes the so-called "viscous displacement." $D_v(t)$ is the viscous displacement at any time t, D_0 is the normal displacement, k is the ratio of the total viscous displacement to the normal displacement and τ is the relaxation time of the aperiodic particles. Pellat's theory implies only one kind of these aperiodic particles all with the same relaxation time. Von Schweidler's extension assumes a large number of different kinds of particles having different relaxation times. The relaxation function then consists of a summation of exponential terms, there being one term for each different relaxation time so that

$$\Phi(t) = \sum_1^n \frac{k_n}{\tau_n} e^{-t/\tau_n}.$$

k_n is proportional to the number of aperiodic particles per unit volume having a relaxation time τ_n. If there are an infinite number of such terms, with relaxation times varying continuously from 0 to ∞, the summation may then be replaced by a definite integral,

$$\Phi(t) = \int_0^\infty \frac{k(\tau)}{\tau} e^{-t/\tau} d\tau, \tag{8}$$

in which $k(\tau)$ may be regarded as the distribution function of the relaxation times.

Substituting Eq. (8) in Eqs. (1), (5) and (6)

$$i_a = C_\infty E \int_0^\infty \frac{k(\tau)}{\tau} e^{-t/\tau} d\tau,$$

$$\epsilon'' = \epsilon_\infty \int_0^\infty d\tau \int_0^\infty \frac{k(\tau)}{\tau} e^{-t/\tau} \sin \omega t \, dt, \tag{1b}$$

and $$\epsilon' = \epsilon_\infty \left[1 + \int_0^\infty d\tau \int_0^\infty \frac{k(\tau)}{\tau} e^{-t/\tau} \cos \omega t \, dt \right].$$

$(k(\tau))/\tau$ is independent of t, hence

$$\epsilon'' = \epsilon_\infty \int_0^\infty \frac{k(\tau)\omega\tau d\tau}{1+\omega^2\tau^2} \tag{5b}$$

and

$$\epsilon' = \epsilon_\infty \left[1 + \int_0^\infty \frac{k(\tau) d\tau}{1+\omega^2\tau^2} \right]. \tag{6b}$$

Although von Schweidler derived equations of the form of (1b), (5b) and (6b) involving a distribution function, he proposed no definite function for the distribution of relaxation times. Instead, he merely used the concept of a distribution of relaxation times as justification for using the empirical relaxation function

$$\Phi(t) = Bt^{-n} \tag{9}$$

in deriving equations for the time variation of the absorption current and the frequency variation of ϵ' and ϵ''. In fact, he equated the theoretical relaxation function containing the distribution function to the empirical relaxation function and determined what the distribution function must be in order that the theoretical and empirical relaxation functions be equivalent. He found that the distribution function must be

$$k(\tau) = \text{const. } \tau^{2-n}.$$

However, it is well known that the empirical relaxation function of Eq. (9) is but an approximation formula which applies only in certain instances over a limited time interval. The introduction of an appropriate distribution function in Eq. (8) should lead to a more general relaxation function. Thus, K. W. Wagner[3] in 1913 proposed that the distribution of relaxation times is governed by a probability function

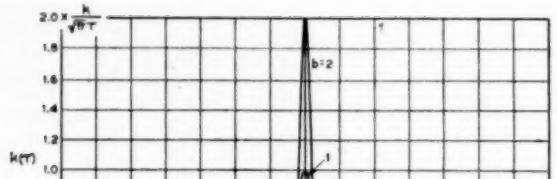


FIG. 1. The effect of the distribution constant b upon the density of distribution of the relaxation times according to the relation $k(\tau)=kb/((\pi)^{\frac{1}{2}}\tau)e^{-(b\ln\tau/\tau_0)^2}$.

analogous to the one proposed by Wiechert[19] for the explanation of the phenomenon of elastic after-effect in solids. The distribution function proposed by Wagner is

$$k(\tau)d\tau=(kb)/(\sqrt{\pi})e^{-b^2z^2}dz, \tag{10}$$

where $z=\ln\tau/\tau_0$.

This law states that the logarithms of the relaxation times of the various terms are grouped about the logarithm of the most prominent relaxation time τ_0, the density of distribution being governed by the distribution constant b as illustrated in Fig. 1. When b is ≥ 2, the grouping of relaxation times about the most prominent value τ_0 is very dense. As b decreases from 2 to 0, the density of distribution broadens rapidly becoming infinitely broad as $b\to 0$.

Introducing this probability function into Eqs. (8), (1b), (5b) and (6b)

$$\Phi(t)=\frac{kb}{\sqrt{\pi}\tau_0}\int_{-\infty}^{\infty}e^{-b^2z^2-z-(t/\tau_0)e^{-z}}dz, \tag{8a}$$

$$i_n=\frac{C_\infty Ekb}{\sqrt{\pi}\tau_0}\int_{-\infty}^{\infty}e^{-b^2z^2-z-(t/\tau_0)e^{-z}}dz, \tag{1c}$$

$$\epsilon''=\frac{\epsilon_\infty kb}{\sqrt{\pi}}e^{-b^2z_0^2}\int_0^{\infty}e^{-b^2u^2}\frac{\cosh 2b^2z_0u}{\cosh u}du,^{20} \tag{5c}$$

and

$$\epsilon'=\epsilon_\infty\left[1+\frac{kb}{\sqrt{\pi}}e^{-b^2z_0^2}\int_0^{\infty}e^{-b^2u^2}\times\frac{\cosh(2b^2z_0-1)u}{\cosh u}du\right],^{20} \tag{6c}$$

where $z_0=\ln\omega\tau_0$,

$z=\ln\tau/\tau_0$,

and $u=z+z_0=\ln\omega\tau$.

These are the equations derived by Wagner for the relaxation function, the time variation of the reversible absorption current, and the frequency variation of ϵ'' and ϵ', respectively. The definite integrals of these equations cannot be evaluated in closed form and hence must be evaluated numerically by approximation methods. Wagner[21] has prepared a table of values for the integral of Eq. (1c) for values of b and t/τ_0 ranging from 0 to 1 and 0 to 25, respectively, and has also indicated a method[3] of evaluating the integrals of Eqs. (5c) and (6c) for values of $b\geq 1$.

The development of the Wagner equation for the time variation of the reversible absorption current has been included in this paper for the sake of completeness. Hereafter, the discussion will be limited to the frequency variation of ϵ'' and ϵ' as given by Eqs. (5c) and (6c). It is of interest to note, however, that the same four constants appear in all three equations. Hence, if these constants can be evaluated from experimental data on the time variation of the reversible absorption current, the corresponding frequency variation of ϵ' and ϵ'' may be determined and *vice versa*.

CONSEQUENCES OF WAGNER'S EQUATIONS

The theoretical values of $\epsilon''/(\epsilon_\infty k)$ and $(\epsilon'-\epsilon_\infty)/(\epsilon_\infty k)$ as computed from a numerical evaluation of the integrals of Eqs. (5c) and (6c) for various values of b (0.08 to 2.0) and z_0 (-6 to $+6$) are presented in Tables I and II and their frequency dependence (in terms of z_0) is shown graphically in Figs. 2 and 3, respectively. When $b=\infty$, Wagner's equations transform into Eqs. (5a) and (6a) for a single relaxation time. In

[19] E. Wiechert, Wied. Ann. **50**, 335 (1893).

[20] A review of Wagner's paper appears in "A Critical Résumé of Recent Work on Dielectrics" by L. Hartshorn, published in J.I.E.E. **64**, 1152 (1926). In this review, cos is used in these equations instead of cosh presumably due to an incorrect translation of the German notation.

[21] K. W. Wagner, E.T.Z. **45**, 1279 (1913).

this case, the frequency band in which dielectric absorption occurs is approximately two decades and ϵ_m'' is equal to $0.5\epsilon_\infty k$. However, as pointed out in the introduction, the frequency band in which dielectric absorption is observed in typical dielectrics is frequently much greater than two decades and ϵ_m'' is frequently much less than $0.5\epsilon_\infty k$. That both of these apparent discrepancies can be accounted for by Wagner's equations can readily be seen by an examination of Figs. 2 and 3 which show clearly the effect of a distribution of relaxation times upon the frequency variation of ϵ'' and ϵ', respectively. One effect of a distribution of relaxation times is to increase the frequency band in which dielectric loss and anomalous dispersion occur, the band becoming wider the broader the density of distribution, that is, the lower the value of b. In fact, as b becomes vanishingly small, ϵ' takes the constant value

$$\epsilon' = \epsilon_\infty(1+0.5k)$$

essentially independent of frequency. A second fundamental effect of a distribution of relaxation times is to progressively decrease the relative magnitude of ϵ_m'' from $0.5\epsilon_\infty k$ for a single relaxation time ($b=\infty$) to zero for an infinitely broad distribution ($b=0$). Thus, when $b=0.25$, ϵ_m'' is approximately equal to $0.2\epsilon_\infty k$ compared to $0.5\epsilon_\infty k$

TABLE I. *Values of $\epsilon''/(\epsilon_\infty k)$ as computed from Wagner's equations.*

z_0	$\omega\tau_0$*		b							
			0.15	0.25	0.35	0.6	0.8	1.0	2.0	∞
0	1.00	1.00	0.1265	0.1938	0.2525	0.3443	0.3882	0.420	0.475	0.5000
±0.5	0.607	1.65	0.1258	0.1919	0.2468	0.3298	0.3660	0.389	0.426	0.4432
±1.0	0.368	2.72	0.1236	0.1862	0.2312	0.2909	0.3101	0.318	0.328	0.3238
±1.5	0.223	4.48	0.1208	0.1750	0.2073	0.2368	0.2388	0.234	0.227	0.2123
±2.0	0.135	7.39	0.1165	0.1601	0.1801	0.1793	0.1706	0.160	0.142	0.1325
±2.5	0.082	12.18	0.1109	0.1426	0.1509	0.1285	0.1128	0.102	0.086	0.0814
±3.0	0.0498	20.09	0.1052	0.1249	0.1195	0.0886	0.0713	0.064	0.053	0.0498
±4.0	0.0183	54.60	0.09115	0.08889	0.0684	0.03521	0.02689	0.02346	0.0195	0.0183
±5.0	0.00674	148.4	0.07590	0.05739	0.0343	0.01338	0.00995	0.00865	0.00710	0.00674
±6.0	0.00248	403.4	0.06065	0.03385	0.0154	0.00496	0.00366	0.00318	0.00263	0.00248

* $\omega\tau_0 = e^{z_0}$.

TABLE II. *Values of $(\epsilon'-\epsilon_\infty)/(\epsilon_\infty k)$ as computed from Wagner's equations.*

z_0	$\omega\tau_0$	b								
		0.08	0.15	0.25	0.35	0.6	0.8	1.0	2.0	∞
−6.0	0.00248	0.752	0.894	0.978	0.9965	0.999903	0.999971	0.999983	0.999992	0.999994
−5.0	0.00674	0.714	0.851	0.954	0.9879	0.99932	0.99978	0.99988	0.999942	0.999955
−4.0	0.0183	0.675	0.798	0.911	0.9646	0.9957	0.99845	0.99909	0.99957	0.99966
−3.0	0.0498	0.632	0.734	0.843	0.910	0.967	0.986	0.9936	0.9968	0.99753
−2.5	0.082	0.611	0.701	0.799	0.874	0.947	0.964	0.985	0.992	0.9933
−2.0	0.135	0.589	0.662	0.749	0.820	0.910	0.946	0.959	0.976	0.9821
−1.5	0.223	0.566	0.623	0.692	0.750	0.839	0.887	0.906	0.942	0.9525
−1.0	0.368	0.544	0.585	0.634	0.671	0.744	0.788	0.816	0.866	0.8808
−0.5	0.607	0.522	0.542	0.566	0.590	0.629	0.658	0.674	0.712	0.7391
0	1.00	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.5000
+0.5	1.65	0.478	0.458	0.435	0.410	0.372	0.342	0.326	0.288	0.2689
+1.0	2.72	0.455	0.415	0.366	0.329	0.256	0.212	0.184	0.134	0.1192
+1.5	4.48	0.433	0.377	0.309	0.250	0.161	0.113	0.094	0.058	0.0475
+2.0	7.39	0.410	0.338	0.251	0.180	0.090	0.054	0.041	0.024	0.0179
+2.5	12.18	0.388	0.299	0.199	0.126	0.053	0.036	0.015	0.008	0.0067
+3.0	20.09	0.367	0.267	0.155	0.091	0.032	0.014	0.0064	0.0032	0.00247
+4.0	54.60	0.325	0.202	0.089	0.0354	0.00430	0.00155	0.00091	0.00043	0.00034
+5.0	148.4	0.286	0.149	0.046	0.0121	0.00068	0.00022	0.00012	0.000058	0.000045
+6.0	403.4	0.248	0.106	0.022	0.0035	0.000097	0.000029	0.000017	0.000008	0.000006

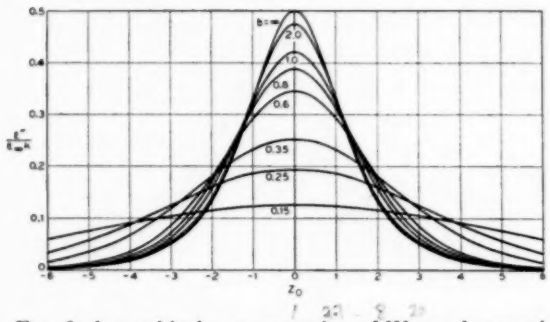


FIG. 2. A graphical representation of Wagner's equation illustrating the effect of the distribution constant b upon the frequency variation of $\epsilon''/(\epsilon_\infty k)\cdot(z_0=\ln \omega\tau_0)$.

when $b=\infty$. For very small values of b, ϵ'' like ϵ' tends to become constant essentially independent of frequency. From Eq. (5c), it can be shown that

$$\lim_{b\to 0} \epsilon''=0.886\epsilon_\infty kb.$$

The width of the frequency band in which anomalous dispersion and dielectric loss are observed in a dielectric and the ratio of ϵ_m'' to $0.5\epsilon_\infty k$ are, therefore, indicative of the density of distribution of the relaxation times in that dielectric.

Other important consequences of Wagner's equations are the following:

I. When $\omega=0$,

$$\left.\begin{aligned}\epsilon_0&=\epsilon_\infty(1+k)\\ \epsilon_0''&=0\end{aligned}\right\}\text{for all values of } b. \qquad (11)$$

II. At the frequency f_m, at which ϵ'' is a maximum,

$$2\pi f_m\tau_0=1, \qquad (12)$$

$$\left.\begin{aligned}\epsilon_m'&=\epsilon_\infty(1+0.5k)\\ \epsilon_m''&=\epsilon_\infty k\cdot\Phi(b).\end{aligned}\right\}\text{for all values of } b, \qquad (13)$$

III. When $\omega=\infty$,

$$\left.\begin{aligned}\epsilon_\infty&=\epsilon_\infty\\ \epsilon_\infty''&=0\end{aligned}\right\}\text{for all values of } b.$$

From Eq. (11), it is seen that

$$\epsilon_\infty k=\epsilon_0-\epsilon_\infty \text{ for all values of } b.$$

$\epsilon_\infty k$ is, therefore, a measure of the total dispersion. Furthermore, the area under the ϵ'' *vs.* $\ln f$ curve for any value of b is likewise a measure of the total dispersion. Sommerman[13] has shown that this area is equal to $(\pi/2)\epsilon_\infty k$ regardless of the form of the distribution function $k(\tau)$.

Graphical method of evaluating the constants of Wagner's equations

In order to compare the frequency variation of ϵ' and ϵ'' as given by Wagner's equations with that actually observed in typical dielectrics in a region of dielectric absorption, the constants of Wagner's equations, namely b, τ_0, ϵ_∞ and k must first be evaluated from the experimental data. Frequently τ_0 and/or ϵ_∞ may be determined directly from the experimental data but often all four constants must be evaluated by a graphical method.

In the method employed in this investigation, a new variable δ is introduced and defined as the ratio of ϵ_a'' at a frequency f_a to ϵ_x'' at a frequency f_x where f_x is any arbitrarily chosen reference frequency and f_a is arbitrarily chosen to be one decade lower than f_x when $f_x<f_m$ and one decade higher than f_x when $f_x>f_m$. Under these conditions, δ will never be greater than one and will have the same value for $\pm z_x$, where

$$z_x=\ln (f_x/f_m). \qquad (14)$$

For a given fixed frequency ratio, δ is independent of ϵ_∞ and k, but is a function of b which governs the slope of the loss-frequency curve and of z_x which determines the position of f_x on the loss-frequency curve with respect to f_m. The theoretical dependence of δ upon b for various values of z_x for

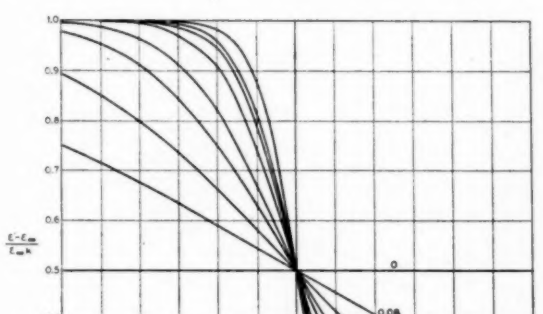


FIG. 3. A graphical representation of Wagner's equation illustrating the effect of the distribution constant b upon the frequency variation of $(\epsilon'-\epsilon_\infty)/(\epsilon_\infty k)\cdot(z_0=\ln \omega\tau_0)$.

TABLE III. *The variation of δ with b and z_x.*

	b								
z_x	0	0.15	0.25	0.35	0.6	0.8	1.0	2.0	∞
0	1	0.903	0.774	0.645	0.4300	0.3452	0.2940	0.224	0.1980
±0.25	1	0.876	0.729	0.589	0.367	0.2830	0.2360	0.1782	0.1596
±0.5	1	0.854	0.688	0.535	0.3105	0.2336	0.1992	0.1495	0.1363
±1.0	1	0.817	0.612	0.441	0.2250	0.1708	0.1504	0.1228	0.1134
±1.5	1	0.778	0.548	0.371	0.1775	0.1382	0.1238	0.1100	0.1049
±2.0	1	0.742	0.494	0.311	0.1505	0.1242	0.1125	0.1056	0.1018
±3.0	1	0.673	0.387	0.233	0.1135	0.0998	0.0935	0.088	0.1002
±4.0	1	0.628	0.319	0.193	0.105	—	—	—	0.1000

a frequency ratio of one decade is given in Table III and presented in the form of reference curves in Fig. 4.

When f_m can be determined directly from the experimental data, b may be determined directly from the reference curves by computing δ at any convenient value of z_x, e.g., $z_x=0$. When the measured frequency band does not include f_m, both b and f_m must be determined simultaneously from two values of δ, say δ_1 and δ_2, computed at reference frequencies f_{x1} and f_{x2}. The values of b and z_{x1} giving δ_1 and the values of b and z_{x2} giving δ_2 are then determined from the reference curves. However,

$$z_{x1}=z_{x2}+\ln(f_{x1}/f_{x2}).$$

Hence b may be plotted as a function of z_{x1}. This plot will consist of two curves the intersection of which determines b and z_{x1}. f_m and τ_0 may then be evaluated by Eqs. (14) and (12), respectively. A representative plot from the analysis on Vinsol at 98°C is shown in Fig. 5. Two reference frequencies and hence two curves are sufficient to determine b and z_{x1}, but a third reference frequency and curve is included as a check.

Having evaluated b and τ_0, the appropriate theoretical values of $\epsilon''/(\epsilon_\infty k)$ and $(\epsilon'-\epsilon_\infty)/(\epsilon_\infty k)$ at the frequencies at which measurements have been made may now be obtained from a plot of the theoretical frequency variation of $\epsilon''/(\epsilon_\infty k)$ and of $(\epsilon'-\epsilon_\infty)/(\epsilon_\infty k)$ for the appropriate values of b and τ_0. The interpolation curves of Figs. 6 and 7 make it possible to construct such plots for any value of $b\leq 2$. If Wagner's equations explain the observed data correctly, then a plot of the observed dielectric loss factor, $\epsilon_{obs.}''$, at various frequencies as a function of $\epsilon''/(\epsilon_\infty k)$ at corresponding frequencies should be a straight line passing through the origin with slope equal to $\epsilon_\infty k$. Similarly, a plot of the observed dielectric constant, $\epsilon_{obs.}'$, at various frequencies as a function of $(\epsilon'-\epsilon_\infty)/(\epsilon_\infty k)$ at corresponding frequencies should be a straight line whose slope is also $\epsilon_\infty k$ and intercept on the ϵ' axis is ϵ_∞. The accuracy with which the points define straight lines is a measure of the accuracy with which Wagner's equations explain the observed data. A representative plot from the analysis on Vinsol at 98°C is shown in Fig. 8. If ϵ_∞ and ϵ_m' can be determined by observation from the experimental curve, k can be evaluated directly by Eq. (13). If ϵ_∞ is unknown but ϵ_0 and ϵ_m' are known, then both ϵ_∞ and k can be evaluated by solving Eqs. (11) and (13) simultaneously.

Having evaluated b, τ_0, ϵ_∞ and k, ϵ_w'' and ϵ_w' can be calculated at any frequency from the corresponding values of $\epsilon''/(\epsilon_\infty k)$ and $(\epsilon'-\epsilon_\infty)/(\epsilon_\infty k)$, respectively. Using the same values of ϵ_∞, k and τ_0, ϵ_s'' and ϵ_s' may be calculated at any frequency from Eqs. (5a) and (6a) for a single relaxation time. The observed and calculated values of ϵ'' and ϵ' may then be compared. In order to illustrate the method described above, a step-by-step evaluation of the constants for Vinsol at 98°C is presented in Table IV.

THE DISTRIBUTION OF RELAXATION TIMES IN TYPICAL DIELECTRICS

As pointed out previously, the consensus of opinion of both the early and recent investigators is that such dielectrics for which the simple absorption equations fail to explain the

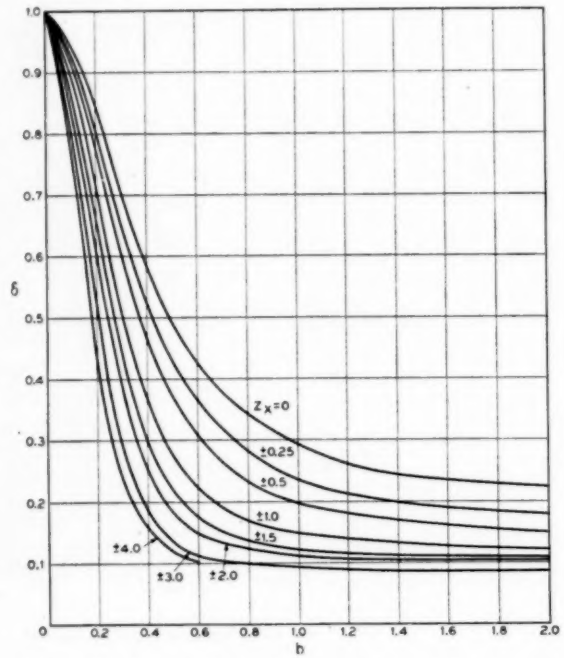


FIG. 4. Reference curves for determining values of b and the corresponding values of z_x giving any value of δ computed from experimental data.

experimental data are polydisperse systems. Wagner's equations are based upon the hypothesis that polydisperse systems are characterized by a statistical distribution of relaxation times. Therefore, if Wagner's equations explain quantitatively the experimental data on a given dielectric, the hypothesis of a statistical distribution of relaxation times is established for that dielectric and the density of the distribution may be determined.

Before analyzing the data for various dielectrics on the basis of Wagner's hypothesis, it is advisable to consider qualitatively the various factors which can give rise to a distribution of relaxation times in a dielectric. Although Wagner's equations apply for any mechanism of dielectric absorption, the factor or factors giving rise to the distribution of relaxation times can be entirely different. Thus, according to the Maxwell-Wagner theory of interfacial polarization, dielectric absorption occurs in some frequency band when a material of properties ϵ_2, σ_2 is distributed in concentration p in a medium of properties ϵ_1, σ_1 provided $\epsilon_2\sigma_1 \neq \epsilon_1\sigma_2$. If component 2 is distributed uniformly in the form of spheres in component 1 and if furthermore $p \ll 1$, then the relaxation time is given by

$$\tau = (2\epsilon_1 + \epsilon_2)/(4\pi(2\sigma_1 + \sigma_2)). \tag{15}$$

Now, if component 2 is considered as being not pure material 2 but rather a mixture of material 2 in material 1, then ϵ_2 and σ_2 will depend upon the composition of the mixture. If there are n kinds of spheres having n different compositions and if, furthermore, there is a distribution of composition about a most prominent value, there must necessarily also be a distribution of relaxation times. A perfectly uniform mixture would have a singular relaxation time but a nonuniform mixture would have a distribution of relaxation times the density of distribution being governed by the degree of nonuniformity. Wagner's paper[17] should be consulted for a more detailed discussion of the distribution of relaxation times in heterogeneous mixtures.

According to the Debye theory of polar molecules, the relaxation time of a polar dielectric is

$$\tau = ((\epsilon_0 + 2)/(\epsilon_\infty + 2))\tau', \tag{16}$$

where

$$\tau' = \zeta/(2kT). \tag{17}$$

τ' is the relaxation time of the polar molecules or aggregates, k is Boltzmann's constant, T is the absolute temperature and ζ is a constant meas-

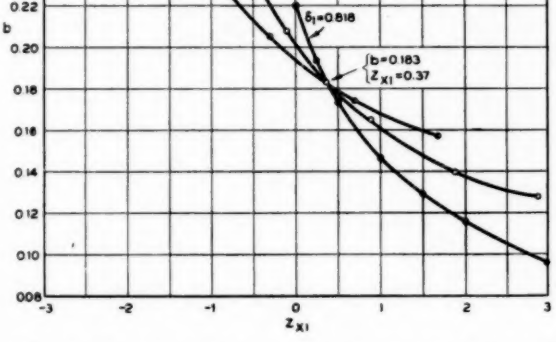


FIG. 5. The graphical determination of the distribution constant b and the most probable relaxation time τ_0 for Vinsol at 98°C.

uring the internal friction. In considering the factors which can give rise to a distribution of relaxation times in a polar dielectric, it will be assumed that the distribution of τ is equivalent to a distribution of τ'.[22] From Eq. (17) it is obvious that a distribution of τ' in a polar dielectric corresponds to a distribution of ζ. According to Stokes' law, the internal friction opposing the rotation of a sphere of radius a in a liquid of inner viscosity η is

$$\zeta = 8\pi\eta a^3.$$

Although this relation is valid only for a spherical particle, it seems reasonable to conclude that, whatever the shape of a particle, ζ will always be a function of its size and shape and of the internal viscosity opposing its rotation. Consequently, the distribution of ζ in a polar dielectric is governed by the variations in these three factors.

Colloidal, associated and polymerized systems are composed of particles of various sizes and perhaps various shapes. The inner viscosity opposing the rotation of a particle is probably a function of its size, shape and position with respect to neighboring particles as pointed out by Sommerman[13] and Oncley and Williams.[11] Sommerman regards the inner viscosity acting upon a spherical particle as being a function of the ratio of the particle radius to the intermolecular distances of the medium. In solutions

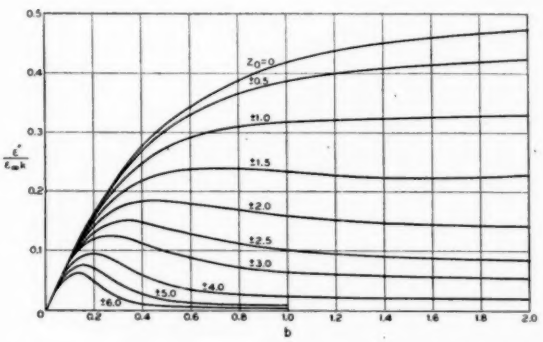


FIG. 6. Interpolation curves for determining the frequency variation of $\epsilon''/(\epsilon_\infty k)$ for any value of $b \leq 2$. $z_0 = \ln \omega\tau_0$.

[22] ϵ_0 in Eq. (16) may be a function of τ' so that the function for the distribution of τ' may not have the same form as that for the distribution of τ. However, the form of the distribution function is unimportant in considering the factors giving rise to the distribution.

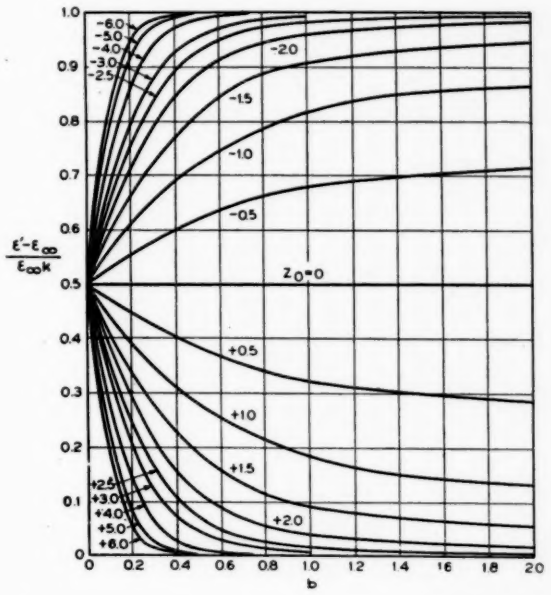


FIG. 7. Interpolation curves for determining the frequency variation of $(\epsilon'-\epsilon_\infty)/(\epsilon_\infty k)$ for any value of $b \leq 2$. $z_0 = \ln \omega\tau_0$.

the intermolecular distances supposedly depend upon the solvent and also vary within it. The inner viscosity assumes the experimental or macroscopic value of viscosity for a large sphere but assumes a lower value for an ion or small polar molecule. Oncley and Williams found that a change in the macroscopic viscosity of floricin solutions by a factor of several thousand apparently does not change the inner frictional constant ζ by any large amount. In order to account for this behavior, they offer the following explanation: "Liquids of high viscosity may contain large molecules but they will also contain some smaller ones. Thus, a viscous mineral oil consists of long chain solid hydrocarbons dissolved in shorter chain liquid hydrocarbons. In a system consisting of polar molecules dissolved in such a solvent, we may expect the arrangement of the polar molecules to be such that the frictional forces would be as small as possible. Thus, the molecules will tend to arrange themselves in such a way that the inner frictional constant ζ will be a minimum. If very large molecules are present, they may form pockets in which the smaller molecules would have nearly free rotational movement. Thus, a small inner frictional constant ζ will result from any observations designed to measure it. Very small

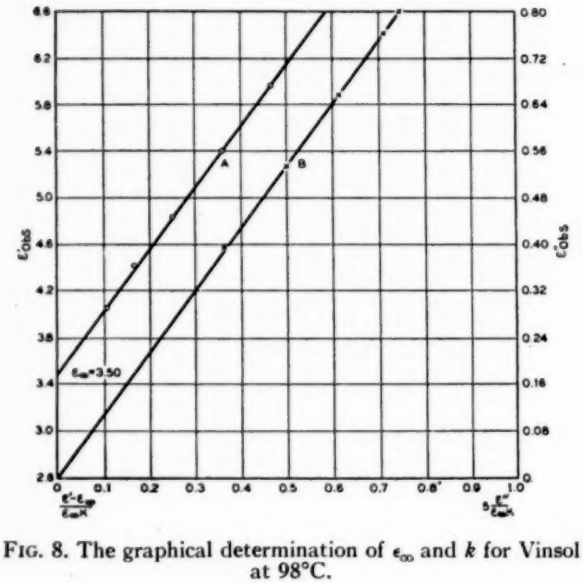


FIG. 8. The graphical determination of ϵ_∞ and k for Vinsol at 98°C.

molecules will also have a small inner frictional constant. Accordingly, an increase in macroscopic viscosity might either increase or decrease the inner frictional constant. It may also be expected that a variable time of relaxation τ, or inner frictional constant ζ, might result even though the polar molecules were only of one kind, the difference being due to their movements with respect to solvent molecules of various sizes and shapes."

It is also conceivable that the individual molecules in an aggregate are capable of being displaced from their preferred orientation by the external field and that the internal friction opposing the displacement of any molecule is a function of the position of that molecule in the aggregate. Thus, it might be expected that the internal friction opposing the displacement of a molecule in the center of an aggregate is much greater than that opposing the displacement of a molecule on the surface of the aggregate.

Although the above considerations are only qualitative, they serve to point out the factor or factors which govern the distribution of relaxation times in dielectrics. Furthermore, the density of the distribution of relaxation times in a dielectric is a measure of the complexity of the constitution of the dielectric. Thus, a perfectly uniform mixture or a pure unassociated polar liquid in which the effect of the intermolecular forces is negligible should be characterized by a single relaxation time whereas a nonuniform mixture or an associated or polymerized polar liquid should be characterized by a distribution of relaxation times, the density of distribution becoming broader the greater the degree of nonuniformity of the mixture or the greater the degree of association or polymerization of the liquid. Some conception of the density of the distribution of relaxation times in a dielectric may be gained from knowledge of the complexity of the dielectric and conversely some conception of the complexity of a dielectric may be gained from an experimental determination of the distribution constant, provided, of course, that Wagner's equations are applicable to the system.

In order to test the validity of Wagner's hypothesis of a statistical distribution of relaxation times in polydisperse systems, the following materials were chosen as being representative of dielectrics of various degrees of complexity:

1. Glycerol[23]—distilled.
2. Vinsol[24]—a mixture of highly oxidized and highly polymerized abietic acid and terpenes.
3. Abietic acid[24]—in the form of fine crystals and reputed to be free from mineral matter and the pitches which are usually associated with rosin.
4. Glycol phthalate resin No. 1—data of R. H. Kienle and H. H. Race.[25]
5. Resoglaz+30 percent Arochlor[24] No. 1254—Resoglaz is a polystyrene thermoplastic resin. Arochlor[10(b)] is a mixture of chlorinated diphenyls.

Glycerol was chosen as being representative of a pure but associated polar liquid. The single molecules of an associated liquid are mutually attracted by secondary valence forces and aggregate into complexes containing two or more of the single molecules. In general, the average number of molecules per complex is small so that the distribution of size of the aggregates and the distribution of inner viscosity impeding their rotation are probably relatively dense. Under these conditions, associated liquids should be characterized by a relatively dense distribution of relaxation times. The results of the analysis on glycerol (Table V and Fig. 9) confirm the supposition that the distribution of relaxation times in glycerol is relatively dense and, further-

[23] S. O. Morgan, Trans. Electrochem. Soc. **65**, 185 (1934).
[24] Unpublished data of the author.
[25] R. H. Kienle and H. H. Race, Trans. Electrochem. Soc. **65** (1934).

more, show that Wagner's equations explain the observed data more accurately than the equations for a monodisperse system. At −31.5°C, the agreement between Wagner's equations and the experimental data is within the experimental error and the computation errors inherent in the many interpolations necessary in computing the four constants of Wagner's equations. At −43.6°C, however, both Wagner's equations and the equations for a monodisperse system give values of ϵ'' which at high frequencies are lower than the experimental values. This departure at high frequencies is to be attributed to the asymmetry of the experimental ϵ'' *vs.* $\ln f$ curve due either to an asymmetrical distribution of relaxation times or to the presence of a second region of dielectric absorption at higher frequencies overlapping the first. Assuming that an asymmetrical distribution of relaxation times is manifestation of a nonuniform distribution of complex size, then the form of the experimental ϵ'' *vs.* $\ln f$ curve would require the distribution of the complexes of smaller than average size to be broader than that of the complexes of greater than average size, whereas the reverse condition appears to be more probable. The fact that the geometric dielectric constant, as determined in this analysis, is greater than the refraction value suggests the presence of a second region of absorption at higher frequencies. In this case,

TABLE IV. *Procedure for evaluating the constants of Wagner's equations from data on Vinsol at 98°C.*

1. Experimental data			2. Compute δ from experimental data	
$f(Kc)$	ϵ'	ϵ''	$f_x(kc)$	$\delta=\epsilon_a''/\epsilon_x''$
1	5.98	0.804		
3	5.41	0.761		
10	4.84	0.658	$f_{x2}=3$	$\delta_2=\epsilon_{30}''/\epsilon_3''=0.702$
30	4.42	0.534		
100	4.07	0.395	$f_{x3}=10$	$\delta_3=\epsilon_{100}''/\epsilon_{10}''=0.600$

3. From Fig. 4, record the values of b and the corresponding values of z_{x1}, z_{x2} and z_{x3} giving δ_1, δ_2 and δ_3, respectively.

$\delta_1=0.818$		$\delta_2=0.702$			$\delta_3=0.600$		
b	z_{x1}	b	z_{x2}	z_{x1}[a]	b	z_{x3}	z_{x1}[b]
0.087	4.0	0.128	4.0	2.9	0.157	4.0	1.7
0.096	3.0	0.140	3.0	1.9	0.174	3.0	0.7
0.116	2.0	0.165	2.0	0.9	0.205	2.0	−0.3
0.130	1.5	0.185	1.5	0.4	0.226	1.5	−0.8
0.152	1.0	0.208	1.0	0.1	0.255	1.0	−1.3
0.174	0.5	0.243	0.5	−0.6	0.306	0.5	−1.8
0.194	0.25	0.269	0.25	−0.85	0.342	0.25	−2.05
0.220	0	0.305	0	−1.10	0.388	0	−2.30

[a] $z_{x1}=z_{x2}+\ln(f_{x1}/f_{x2})=z_{x2}-1.10$.
[b] $z_{x1}=z_{x3}+\ln(f_{x1}/f_{x3})=z_{x3}-2.30$.

4. Plot b *vs.* z_{x1} and determine intersection of curves. See Fig. 5.
   Intersection: $b=0.183$; $z_{x1}=0.37$.
5. Determine f_m and τ_0.
   $z_{x1}=\ln(f_{x1}/f_m)=0.37$; $f_m=690$ cycles per second;
   $\tau_0=2.31\times10^{-4}$ seconds.
6. From Figs. (6) and (7) tabulate $\epsilon''/(\epsilon_\infty k)$ and $(\epsilon'-\epsilon_\infty)/(\epsilon_\infty k)$ for various values of z_0 when $b=0.183$. Plot as a function of f and pick off values at 1, 3, 10, 30 and 100 kc.

z_0	f(kc)	$\frac{\epsilon''}{\epsilon_\infty k}$	$\frac{\epsilon'-\epsilon_\infty}{\epsilon_\infty k}$	f(kc)	$\frac{\epsilon''}{\epsilon_\infty k}$	$\frac{\epsilon'-\epsilon_\infty}{\epsilon_\infty k}$
0	0.69	0.1502	0.500	1	0.1492	0.464
1.0	1.88	0.1460	0.398	3	0.1415	0.357
1.5	3.09	0.1410	0.355	10	0.1230	0.250
2.0	5.10	0.1351	0.309	30	0.1001	0.170
2.5	8.40	0.1250	0.265	100	0.073	0.105
3.0	13.85	0.1176	0.225			
4.0	37.6	0.097	0.155			
5.0	102.3	0.073	0.100			

7. Plot $\epsilon''_{obs.}$ *vs.* $\epsilon''/\epsilon_\infty k$ (Fig. 8B), and $\epsilon'_{obs.}$ *vs.* $(\epsilon'-\epsilon_\infty)/\epsilon_\infty k$ (Fig. 8A).
   Slope either curve $=\epsilon_\infty k=5.37$
   Intercept on ϵ' axis (curve A) $=\epsilon_\infty=3.50$
   Hence $k=1.53$
8. Compute ϵ_w'' and ϵ_w' from $\epsilon''/\epsilon_\infty k$ and $(\epsilon'-\epsilon_\infty)/\epsilon_\infty k$, respectively, and ϵ_s'' and ϵ_s' from Eqs. (5a) and (6a), respectively, using values of ϵ_∞, τ_0 and k evaluated; and compare with $\epsilon''_{obs.}$ and $\epsilon'_{obs.}$.

f(kc)	$\epsilon''_{obs.}$	ϵ_w''[c]	ϵ_s''[d]	$\epsilon'_{obs.}$	ϵ_w'[c]	ϵ_s'[d]
1	0.804	0.801	2.51	5.98	5.99	5.23
3	0.761	0.760	1.17	5.41	5.42	3.77
10	0.658	0.660	0.369	4.84	4.84	3.53
30	0.534	0.537	0.123	4.42	4.41	3.50
100	0.395	0.392	0.037	4.07	4.06	3.50

[c] According to Wagner's equations.
[d] According to equations for single relaxation time.

9. Plot the observed and calculated values of ϵ'' and ϵ' as a function of frequency. See Fig. 9.

the equations for the frequency variation of ϵ' and ϵ'' would contain two Wagner integrals, one for each region of absorption. At frequencies in the band in which the two regions overlap, ϵ'' will be the sum of the individual contributions of each region. Two regions of absorption in glycerol could be accounted for by assuming that there are two most probable sizes of the complexes, one being much larger than the other. The smaller most probable size might possibly correspond to a single molecule. The low and high frequency absorptions would then be due to the restrained rotation of the larger complexes and the smaller complexes or single molecules, respectively. The absorption constant of the low frequency absorption in glycerol is much greater than that of the high so that the number of large complexes is probably much greater than the number of small complexes or single molecules unless the moment of the large complexes is much greater than that of the small. Two regions of absorption in glycerol could also be accounted for if the high frequency absorption is regarded as being due to the restrained rotation of the OH groups. However, the existence of a high frequency absorption in glycerol is uncertain. The above discussion is given primarily to point out the usefulness of an analysis of this kind in interpreting dielectric data.

Vinsol is representative of a dielectric of relatively high degree of polymerization. According to the supplier it is a mixture of highly oxidized and highly polymerized abietic acid and terpenes. On cooling from the liquid state, Vinsol forms a glass and exhibits dielectric absorption due presumably to the restrained rotation of polar aggregates of abietic acid and some of the terpenes. Since both the size of the aggregates and the inner viscosity impeding their rotation probably vary between wide limits in a material of such complex constitution, Vinsol should be characterized by a very broad distribution of relaxation times. The results of the analysis on Vinsol is given in Table V and Fig. 10. As predicted, the distribution of relaxation times was found to be very broad ($b=0.18$). At all three temperatures, Wagner's equations explain the experimental data quantitatively whereas the simple absorption equations fail completely. Thus, ϵ'' calculated from the equations for a monodisperse system for 100 kilocycles at 98°, 10 kilocycles at 110° and 1 kilocycle at 129°C is, respectively, 0.094, 3.45 and 0.031 times as large as the experimental value. It is obvious that the application of these equations to polydisperse systems of such complexity is utterly hopeless.

Abietic acid is representative of a dielectric whose constitution is more complex than that of glycerol but less complex than that of Vinsol. Since abietic acid polymerizes to some extent, the distribution of the ηa^3 factor in it is probably broader than in glycerol which merely associates but, on the other hand, its degree of polymerization is probably lower than that of Vinsol. Therefore, the density of the distribution of relaxation times in abietic acid should be intermediate between that in glycerol and Vinsol. The results of the analysis on abietic acid are given in Table V and Fig. 11. As predicted, the distribution of relaxation times was found to be moderately broad ($b=0.388$), being intermediate between that in glycerol and Vinsol. The agreement between Wagner's equations and the experimental data is again quantitative whereas the agreement with the equations for a monodisperse system is very poor.

Kienle and Race are of the opinion that the dielectric absorption which they observed in glycol phthalate resin No. 1 and other alkyd resins is of the Maxwell-Wagner type. They picture the probable structure of such resins as being that of a gel composed of a heterogeneous mixture of macromolecules or micells of different sizes with conducting components held in the structure. The distribution of components in such a system is undoubtedly nonuniform and hence a distribution of relaxation times is to be expected. The analysis of the data on glycol phthalate resin No. 1, given in Table V and Fig. 12, shows that the distribution of relaxation times in this material is moderately broad ($b=0.327$), comparable to that in abietic acid. Wagner's equations again explain the experimental data quantitatively whereas the equations for a monodisperse system are unsatisfactory. The fact that the geometric dielectric constant is considerably greater than the refraction value is additional evidence in support of the view that a Maxwell-Wagner interfacial

TABLE V. *Comparison of the observed and theoretical frequency variation of ϵ'' and ϵ' for various dielectrics*

		$f(kc)$	$\epsilon_{obs.}''$	ϵ_w''*	ϵ_s''**	$\epsilon_{obs.}'$	ϵ_w'*	ϵ_s'**
			A. Glycerol. Temp. −31.5°C					
b	1.16							
$f_{max.}$	135 kc	1		0.38	0.35	54.2	54.2	54.2
τ_0	1.18×10^{-6} sec.	3	1.3	1.28	1.05	54.2	54.2	54.2
ϵ_∞	6.7	10	4.1	4.18	3.50	53.7	53.5	53.9
k	7.09	30	11.1	10.8	10.1	50.5	50.2	52.0
$\epsilon_\infty k$	47.5	100	19.5	19.9	22.7	35.3	35.9	37.3
			Temp. −43.6°C					
b	1.19							
$f_{max.}$	11 kc	1	5.0	5.1	4.27	57.4	56.8	57.3
τ_0	1.45×10^{-5} sec.	3	13.2	12.6	12.1	53.0	52.5	54.4
ϵ_∞	10.2	10	20.8	20.7	23.6	35.2	35.5	36.2
k	4.66	30	15.8	15.2	15.3	18.6	18.3	15.8
$\epsilon_\infty k$	47.5	100	8.1	6.0	5.2	10.1	11.2	10.8
			B. Vinsol. Temp. 98°C					
b	0.183							
$f_{max.}$	0.69 *kc*	1	0.804	0.801	2.52	5.98	5.99	5.23
τ_0	2.31×10^{-4} sec.	3	0.761	0.760	1.18	5.41	5.42	3.76
ϵ_∞	3.50	10	0.658	0.660	0.371	4.84	4.84	3.52
k	1.53	30	0.534	0.537	0.124	4.42	4.41	3.49
$\epsilon_\infty k$	5.37	100	0.395	0.392	0.037	4.07	4.06	3.49
			Temp. 110°C					
b	0.18							
$f_{max.}$	10 kc	1	0.680	0.680	0.534	7.11	7.14	8.64
τ_0	1.59×10^{-5} sec.	3	0.750	0.759	1.48	6.60	6.62	8.25
ϵ_∞	3.30	10	0.782	0.782	2.70	6.00	6.00	6.00
k	1.63	30	0.764	0.763	1.62	5.44	5.42	3.84
$\epsilon_\infty k$	5.39	100	0.674	0.680	0.534	4.85	4.86	3.35
			Temp. 129°C					
b	0.147							
$f_{max.}$	617 kc	1	0.353	0.336	0.011	7.86	7.89	8.45
τ_0	2.58×10^{-7} sec.	3	0.435	0.442	0.030	7.61	7.61	8.45
ϵ_∞	2.23	10	0.551	0.554	0.108	7.23	7.23	8.45
k	2.79	30	0.642	0.642	0.301	6.81	6.80	8.42
$\epsilon_\infty k$	6.22	100	0.722	0.722	0.981	6.28	6.24	8.30
			C. Abietic Acid. Temp. 86.8°C					
b	0.388							
$f_{max.}$	70 kc	1	0.047	0.046	0.012	3.674	3.673	3.690
τ_0	2.27×10^{-6} sec.	3	0.092	0.089	0.036	3.635	3.634	3.688
ϵ_∞	2.842	10	0.158	0.158	0.119	3.545	3.545	3.674
k	0.298	30	0.213	0.213	0.307	3.401	3.399	3.558
$\epsilon_\infty k$	0.848	100	0.225	0.226	0.399	3.201	3.206	3.071
			D. Glycol Phthalate Resin No. 1. Temp. 30°C					
b	0.327							
$f_{max.}$	300 kc	1	1.00	0.22	0.04	16.5	16.6	16.7
τ_0	5.30×10^{-7} sec.	3	0.68	0.58	0.11	16.2	16.4	16.7
ϵ_∞	5.8	10	1.11	1.11	0.36	15.9	15.9	16.7
k	1.88	30	1.76	1.77	1.08	15.0	15.0	16.6
$\epsilon_\infty k$	10.92	100	2.38	2.36	3.27	13.2	13.2	15.6
		300	2.62	2.62	5.46	11.2	11.3	11.3
		1000	2.28	2.32	3.00	9.3	9.1	6.7
			E. Resoglaz+30 percent Arochlor No. 1254. Temp. 25°C					
b	0.135							
$f_{max.}$	19kc	1	0.0397	0.0397	0.0210	2.73	2.73	2.85
τ_0	8.3×10^{-6} sec.	3	0.0438	0.0438	0.0620	2.70	2.71	2.84
ϵ_∞	2.45	10	0.0467	0.0462	0.1656	2.67	2.67	2.77
k	0.164	30	0.0470	0.0463	0.1815	2.63	2.64	2.57
$\epsilon_\infty k$	0.402	100	0.0443	0.0446	0.0739	2.60	2.60	2.46

* Theoretical values according to Wagner's equations.
** Theoretical values according to simple absorption equations.

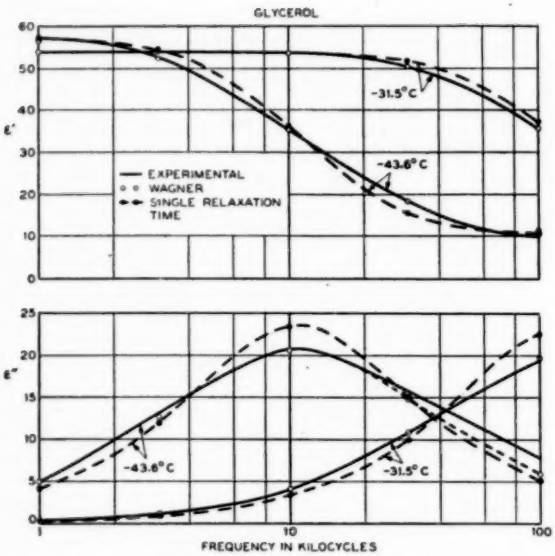


FIG. 9. The experimental and theoretical frequency variation of ϵ' and ϵ'' for glycerol at $-31.5°$ and $-43.6°$C.

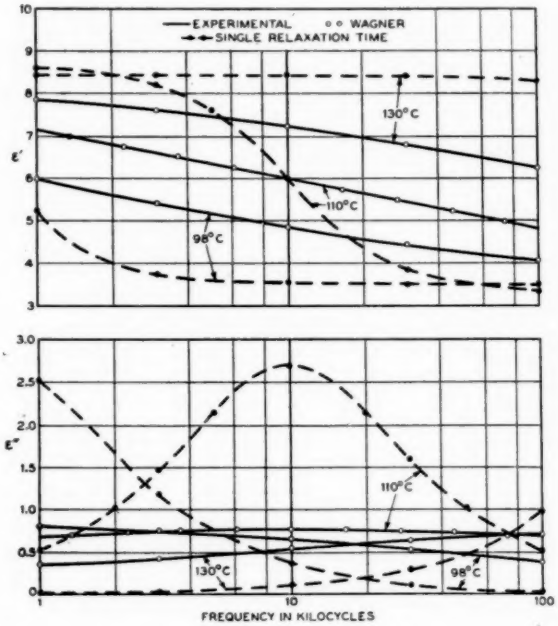


FIG. 10. The experimental and theoretical frequency variation of ϵ' and ϵ'' for Vinsol at three temperatures.

polarization is the cause of the absorption in this material.

The system Resoglaz+30 percent Arochlor (Table V and Fig. 13) is an example of a system in which a polar liquid characterized by a relatively dense distribution of relaxation times is dissolved in a resin characterized by an extremely broad distribution of relaxation times. Resoglaz is very highly polymerized styrene in which the particle size and the inner viscosity impeding the motion of the slightly polar particles probably vary between very wide limits. The fact that both the dielectric constant and dielectric loss factor of Resoglaz are essentially independent of frequency over a wide frequency band, a direct consequence of Wagner's equations as b approaches zero, constitutes experimental evidence of an extremely broad distribution of relaxation times in this material (b is probably of the order of 0.01). Arochlor No. 1254, a mixture of chlorinated diphenyls, exhibits anomalous dispersion and dielectric loss presumably due to the restrained rotation of polar molecules or aggregates. At 25°C, ϵ'' attains its maximum value around 300 kilocycles and the distribution of relaxation times appears to be relatively dense (b=about 1). The absorption observed in the system Resoglaz+30 percent Arochlor No. 1254 is undoubtedly due to the Arochlor but in this system at 25°C, ϵ'' attains its maximum value at 19 kilocycles and the distribution of relaxation times is very broad (b=0.135). The Arochlor molecules or aggregates dissolved in the Resoglaz appear to be acted upon by a greater and more variable internal friction than the molecules in the bulk material.

CONCLUSIONS

The results of these analyses show definitely that the simple absorption equations derived for a monodisperse system are not applicable to dielectrics of a polydisperse nature. Although these simple equations give fair agreement with the experimental data on dielectrics of relatively simple constitution such as associated liquids, the agreement becomes increasingly poor as the complexity of the dielectric increases. For dielectrics of even moderate complexity, these equations are worthless. Wagner's equations, on the other hand, appear to be applicable to polydisperse systems of all degrees of complexity. The quantitative agreement obtained between Wagner's equations and the experimental data on the various dielectrics investigated with the possible exception of glycerol is indeed striking and is viewed as a confirmation of Wagner's hypothesis of a statistical distribution of relaxa-

tion times in these polydisperse systems. If an analysis of a polydisperse system on the basis of Wagner's hypothesis discloses that Wagner's equations are not applicable to the system, the indications are that the distribution function governing the distribution of relaxation times has a different form than the one proposed by Wagner or that there are two or more overlapping regions of absorption or that there are other disturbing factors which must be taken into consideration.

Important information concerning the dielectric behavior of a polydisperse system may be obtained by analyzing, on the basis of Wagner's hypothesis, experimental data on the frequency variation of ϵ' and ϵ'' over a frequency band which need include only a suitable portion of a region of absorption. This analysis will show whether or not Wagner's equations are applicable to the system. If they are, a statistical distribution of relaxation times in the system is indicated and the complete curves for the frequency variation of ϵ' and ϵ'' and the corresponding curve for the time variation of the reversible absorption current can be calculated, provided of course that there are no other regions of absorption in the material. The evaluation of the constants of these equations gives at once the static and geometric dielectric constants, the most probable relaxation time, the distribution constant, the frequency at which ϵ'' is a maximum and the absorption constant of the material, all of which are important properties of a dielectric.

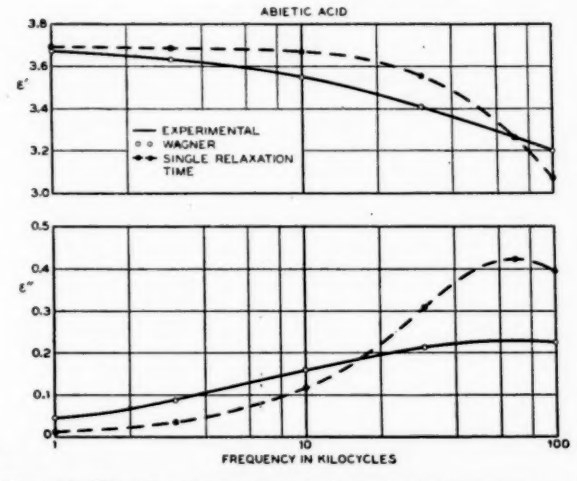


FIG. 11. The experimental and theoretical frequency variation of ϵ' and ϵ'' for abietic acid at 86.8°C.

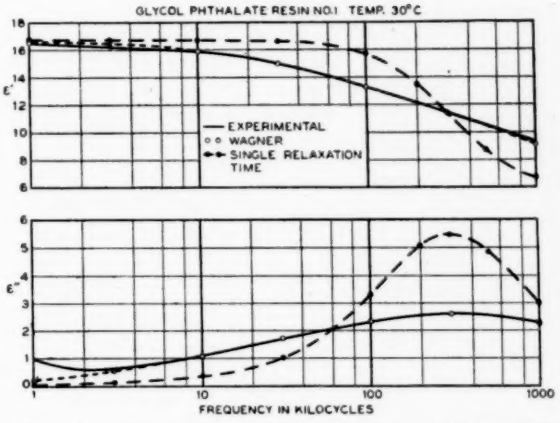


FIG. 12. The experimental and theoretical frequency variation of ϵ' and ϵ'' for glycol phthalate resin No. 1 at 30°C.

The distribution constant is a rough index of the complexity of a dielectric. It may be used as a qualitative measure of the degree of association or polymerization of a dielectric of the polar type and of the degree of uniformity of the distribution of the various components of a heterogeneous mixture of the Maxwell-Wagner type. An estimate of the complexity of a dielectric may therefore be obtained from an experimental determination of its distribution constant. An investigation of the relationship between the distribution constant and the degree of association or polymerization of a dielectric would be

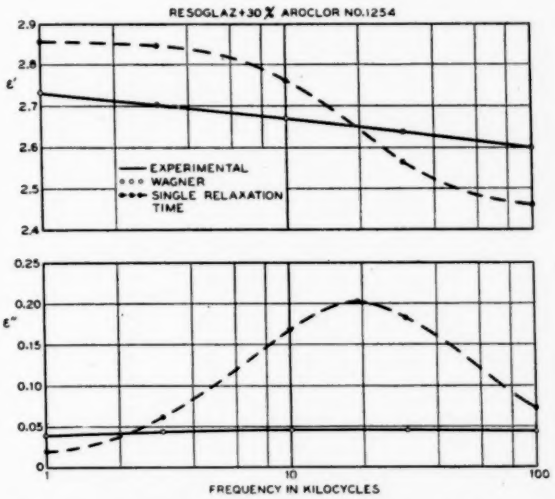


FIG. 13. The experimental and theoretical frequency variation of ϵ' and ϵ'' for Resoglaz+30 percent Arochlor No. 1254 at 25°C.

of interest as it might throw some light on the general problem of association and might provide a useful tool for controling the degree of polymerization in the manufacture of certain resins.

The inherent difficulties in ascertaining whether an observed absorption is of the Debye type or of the Maxwell-Wagner type is well known. Unfortunately, an analysis of experimental data on the basis of Wagner's hypothesis appears to be of little value in differentiating between the two mechanisms of dielectric absorption. Although the factors causing a distribution of relaxation times in the two mechanisms are entirely different, a very dense or a very broad distribution is theoretically possible in either. In some cases, however, the magnitude of ϵ_∞ may offer a clue as to the nature of the absorption. Thus, in systems in which ϵ_∞ is considerably greater than the refraction value, the absorption is more apt to be of the Maxwell-Wagner type, unless there is evidence of a second region of absorption at higher frequencies as in glycerol, whereas in systems in which ϵ_∞ is approximately equal to the refraction value, the absorption is more apt to be of the Debye type.

The author is greatly indebted to Dr. L. A. MacColl of Bell Telephone Laboratories for the numerical evaluations of the complex integrals of Wagner's equations and other mathematical assistance and to Dr. S. O. Morgan, also of Bell Telephone Laboratories, for his interest and helpful suggestions in connection with this investigation.

Electron Optics of an Image Tube

G. A. MORTON AND E. G. RAMBERG, *RCA Manufacturing Company, Inc., Camden, N. J.*

(Received September 3, 1936)

(1). *Introduction:* Description of fixed-focus and variable-focus electrostatic image tube; scope of paper. (2). *Focusing properties:* Measurement of the variation of image distances and magnification with object distance and ratio of applied voltages for both types of tubes; calculation of potential distributions and electron paths; comparison of experimental and theoretical results. (3) *Aberrations:* Classification of aberrations; measurement of tangential and sagittal image surfaces for fixed-focus tube; calculation of axial (chromatic and spherical) aberrations; calculation of field aberrations; comparison between measurements and calculation; reduction of field aberrations by curving the cathode.

1. INTRODUCTION

IN a recent paper, Drs. Zworykin and Morton[1] have described several electron optical systems for producing, on a fluorescent screen, slightly magnified or unmagnified images of extended objects, such as photosensitive surfaces on which light images are projected. The combination of such a photocathode, a suitable electron lens system and a fluorescent screen for observing the electron image, we shall call an "image tube." In the image tubes considered in the paper referred to above, the focusing of the electrons was accomplished by purely electrostatic means. The simplest ("fixed-focus") type of tube (Fig. 1, a) consisted of two narrowly separated cylinders of equal diameter at cathode and screen potential, respectively, the juncture of the two being described as the "lens." By replacing the cathode cylinder by a system of rings at uniformly increasing or decreasing potential (Fig. 1, b), simulating a resistive cylinder between whose ends a voltage is applied, it was found possible to obtain exact focusing for a given object and image distance by varying the ratio of the voltages applied across the system of rings (V_2) and that between the anode and cathode (V_1). Finally, it was found that, in either case, giving the cathode a curvature such that its center of curvature falls into the lens (Fig. 1, c), greatly improves the quality of the image, especially in the peripheral portions.

In the present paper, a more detailed experimental and theoretical study of the focusing properties and image defects of the above image tubes will be presented. It is hoped that such a study may help to outline the scope of these image tubes by pointing out the natural limits imposed by the focusing characteristics and aberrations, and suggesting in what directions future improvement must be sought.

2. FOCUSING PROPERTIES

In order to measure the variation of image distance v with object distance u, both measured from the "lens," in the fixed-focus type of tube, the tube shown in Fig. 2 was constructed. The

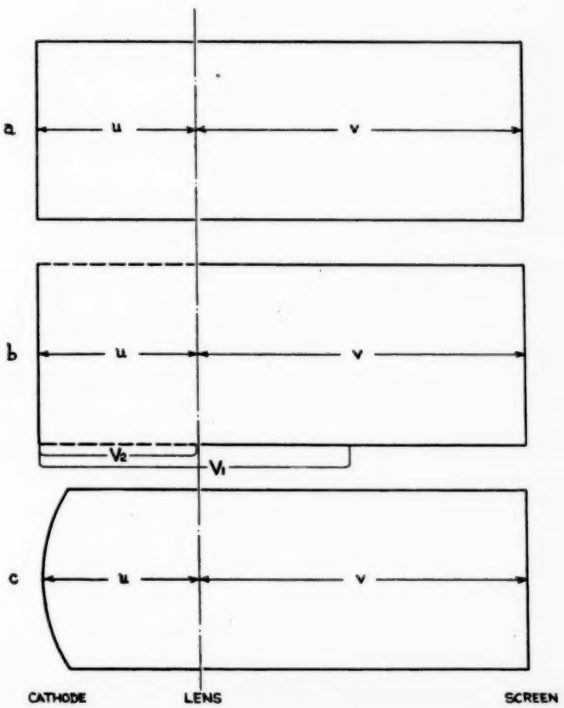


FIG. 1. Schematic diagram of image tubes.

[1] V. K. Zworykin and G. A. Morton, J. Opt. Soc. Am. **26**, 181 (1936).

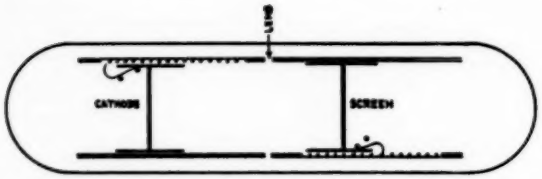


FIG. 2. Experimental tube with movable cathode and screen.

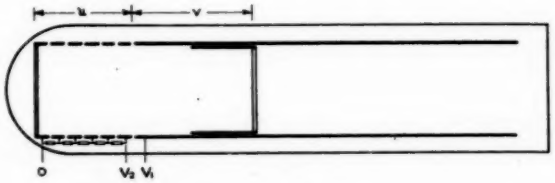


FIG. 3. Variable focus tube with movable screen.

cathode and screen are movable and can be shifted by tilting the tube. Catches are provided which hold the cathode in place when the tube is tilted in one position, and the screen in place when it is rotated about its axis through 180°. This allows the two elements to be moved independently. The lens-to-image distance v, for which the center of the image appeared sharp, as well as the corresponding magnification m, were measured for a set of cathode-to-lens distances u.

Fig. 3 shows a corresponding modification of the focusing ring type tube shown in Fig. 1, b. Here, the screen is made movable, and the cathode fixed at a distance of 1.8 lens radii from the lens, the image distance and magnification being measured for different values of V_2/V_1, the ratio of the voltage applied between the cathode and outermost ring, and that applied between the cathode and anode.

To obtain similar information by calculation, it is necessary first to find the variation of the refractive index n, of the medium, which is conveniently given by the square root of the potential, if we choose the latter to be zero at points where the velocity of the electrons vanishes (e.g., at the cathode if we neglect the initial velocities).[2] Then, with the aid of Fermat's law,

$$\delta \int n ds = 0, \tag{1}$$

ds being an element of path of the ray considered, and the variation applying to paths with fixed endpoints, the paths actually taken by electrons emitted from the cathode and, hence, the position and magnification of the image can be determined.

The variation of potential within the tube is governed by Laplace's equation and conditions imposed at the boundaries, i.e., the potentials applied to the cylinders and the cathode. We shall, for convenience, imagine the anode cylinder to be extended to infinity. This introduces no appreciable error, as electrons which have entered the anode cylinder, owing to their high velocities, are but slightly affected by any "end-effects," including among these the presence of the fluorescent screen.

Cylindrical coordinates, with their origin at the center of the cathode, are most convenient. z will denote the distance from this point along the axis of symmetry, and r the separation of any point from this same axis. The radius of the focusing cylinders ("lens radius") will be taken as unit of length. The substitution

$$\chi(r, z) = F(z) \cdot G(r) \tag{2}$$

in Laplace's equation

$$\nabla^2 \chi = 0 \tag{3}$$

reduces its solution to that of the two ordinary differential equations.

$$(1/F)(d^2F/dz^2) = -k^2, \tag{4a}$$

$$(1/rG)(d/dr)(rdG/dr) = k^2, \tag{4b}$$

k^2 being a separation parameter, with the general solutions

$$F = ae^{ikz} + be^{-ikz}, \tag{5a}$$

$$G = cJ_0(ikr) + dN_0(ikr). \tag{5b}$$

The solution for the potential must have the form of a linear combination of products of such functions with the same value of k, which may be any value in the complex plane. The requirement that the potential remain finite as $z \to \infty$, eliminates terms with complex k. Furthermore, the condition that it be finite on the axis, eliminates those containing the Neumann function N_0. Finally, the fact that the potential vanishes (in the case of the flat cathode) all over the cathode plane, limits the trigonometric functions to be considered to sines. The solution for the potential thus becomes

[2] See e.g., E. Brüche and O. Scherzer, *Geometrische Elektronenoptik* (Springer, 1934), p. 3.

$$\varphi(r, z)=\int_0^\infty b_k J_0(ikr)\sin kz dk. \quad (6)$$

The coefficient b_k can be evaluated by substituting the values of $\varphi(r, z)$ for $r=1$ by the usual methods used for evaluating Fourier coefficients. For the fixed-focus tube, assuming the gap between the cathode and anode cylinders to be negligible, the potential along the axis becomes

$$\Phi(z)=\varphi(0, z)=\frac{2}{\pi}\int_0^\infty V_1\frac{\cos ku}{k}\frac{\sin kz}{J_0(ik)}dk, \quad (7)$$

the first two derivatives being given correspondingly by

$$\Phi'(z)=\frac{2}{\pi}\int_0^\infty V_1\cos ku\frac{\cos kz}{J_0(ik)}dk, \quad (7')$$

$$\Phi''(z)=\frac{-2}{\pi}\int_0^\infty V_1 k\cos ku\frac{\sin kz}{J_0(ik)}dk. \quad (7'')$$

Fig. 4 shows the axial potential distribution of the fixed-focus type of tube with $u=1.8$, together with its first four derivatives.

Similarly, for the focusing ring type tube, replacing the set of rings by a cylinder of uniformly increasing or decreasing potential, the result for the potential along the axis is

$$\Phi(z)=\frac{2}{\pi}\int_0^\infty\left(\frac{V_2}{u}\frac{\sin ku}{k^2}+(V_1-V_2)\frac{\cos ku}{k}\right)\frac{\sin kz}{J_0(ik)}dk. \quad (8)$$

For any given object distance u, it is thus possible, by numerical or graphical quadrature of the above integrals, to find the potential along the axis for different values of z. Furthermore, the potential off the axis is given in terms of the axial potential and its derivatives by the relation[3]

$$\varphi(r, z)=\Phi(z)-\frac{r^2}{4}\Phi''(z)+\frac{r^4}{64}\Phi^{IV}(z)-\cdots. \quad (9)$$

Replacing, temporarily, the cylindrical coordinates r, ϑ by Cartesian coordinates x, y ($r^2=x^2+y^2$), Fermat's law can be written

$$\delta\int F(x, y, x', y', z)dz=\delta\int\{\varphi(x, y, z)\times[x'^2+y'^2+1]\}^{\frac{1}{2}}dz=0, \quad (10)$$

where a prime denotes differentiation with respect to z. The corresponding Euler equations for the electron paths,

$$(d/dz)F_{x'}-F_x=0, \quad (11a)$$

$$(d/dz)F_{y'}-F_y=0, \quad (11b)$$

thus become

$$x''=\frac{x'y'y''}{1+y'^2}-\frac{1}{2\varphi}x'\left(1+\frac{x'^2}{1+y'^2}\right)\times\left(\frac{\partial\varphi}{\partial r}\left[\frac{x'x+y'y}{r}\right]+\frac{\partial\varphi}{\partial z}\right)+\frac{(1+x'^2+y'^2)^2}{1+y'^2}\frac{x}{r}\frac{1}{2\varphi}\frac{\partial\varphi}{\partial r}, \quad (12a)$$

$$y''=\frac{x'y'x''}{1+x'^2}-\frac{1}{2\varphi}y'\left(1+\frac{y'^2}{1+x'^2}\right)\times\left(\frac{\partial\varphi}{\partial r}\left[\frac{x'x+y'y}{r}\right]+\frac{\partial\varphi}{\partial z}\right)+\left(\frac{1+x'^2+y'^2}{1+x'^2}\right)^2\frac{y}{r}\frac{1}{2\varphi}\frac{\partial\varphi}{\partial r} \quad (12b)$$

When the rays considered are in meridian planes only, attention may be confined to the first of these equations, and x may be replaced throughout by r, y being set equal to zero. If,

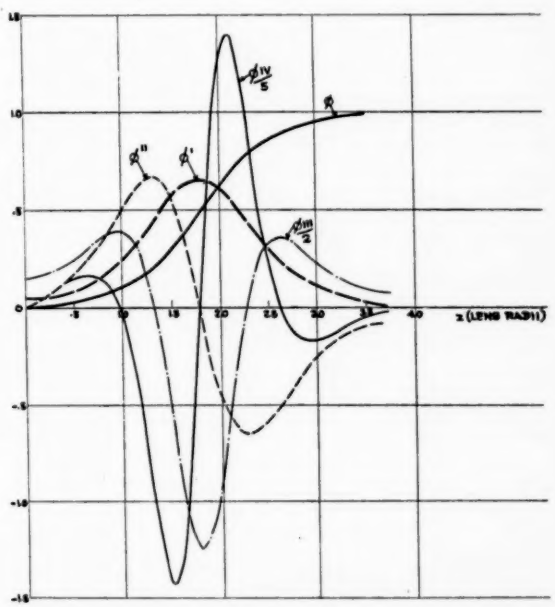


FIG. 4. Axial potential for fixed-focus tube and its first four derivatives.

[3] Reference 2, p. 66, Eq. (16).

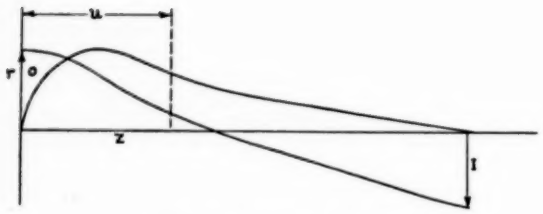


FIG. 5. Electron paths.

finally, the expression (9) is substituted for φ we may drop, for "paraxial rays," i.e., such that deviate infinitesimally from the axis of the system, all terms involving higher powers than the first of r and r' yielding:

$$r''=-(\Phi'/2\Phi)r'-(\Phi''/4\Phi)r. \qquad (13)$$

If the radial distance from the center of the cathode to the starting point of an electron ray, $r(0)$, and its initial slope, $r'(0)$, are given, this equation, by a process of numerical integration, enables us to find the distance of the ray from the axis, $r(z)$, for any value of z.

To locate the image position, we shall consider the path of an electron leaving the center of the cathode with an infinitesimal initial radial velocity. The determination of the path is simplified by introducing, as dependent variable,

$$b=1/2z-1/r(dr/dz). \qquad (14)$$

which reduces Eq. (13) to a first-order equation. The term $-(dr/dz)/r$ is the convergence of the ray at the point z, i.e., the reciprocal of the segment of the axis measured from the abscissa of the point considered to the intersection with the tangent to the ray; $-1/2z$, on the other hand, is the initial convergence of the ray, which describes a parabola of the form $r^2=pz$ ($p=a$ constant) in the uniform field immediately in front of the cathode. The convergence variable, b, is thus chosen so as to vanish for $z=0$, where the convergence itself becomes negatively infinite.

Substituting (14) in (13), the equation governing b becomes

$$\frac{db}{dz}=b^2-b\left(\frac{1}{z}+\frac{\Phi'}{2\Phi}\right)+\frac{\Phi''}{4\Phi}+\frac{1}{2z}\left(\frac{\Phi'}{2\Phi}-\frac{1}{2z}\right). \qquad (15)$$

This equation, again, cannot be integrated analytically but must be evaluated numerically. The value of b is determined, by a step-by-step process of integration, for some particular value of z, far enough to the right of the lens so that the ray path is essentially a straight line. From the convergence of the ray at this point the image position is located.

To determine the magnification, as illustrated in Fig. 5, the point of intersection of the image plane, as determined above, and the path of an electron leaving the cathode, with zero initial velocity, at a radial distance "O" from the center of the cathode, is located. The radial distance "I" of this intersection, divided by the distance "O," gives the magnification. To calculate this path Eq. (13) must be used.

A comparison of the results of the measurements and calculations for the fixed and variable focus type tubes is given by Figs. 6 to 9. The relatively large deviations for very small image distances, v, may be ascribed to the interference of the screen with the field distribution within the lens. The closeness with which the magnification is given by $v/2u$ is striking for both types of tubes.

3. ABERRATIONS

The aberrations which detract from the quality of the electron image observed on the fluorescent

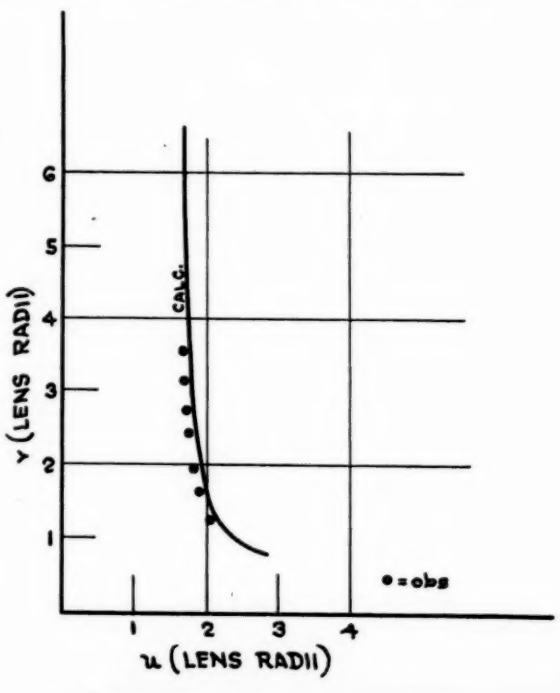


FIG. 6. Focusing relation (fixed focus).

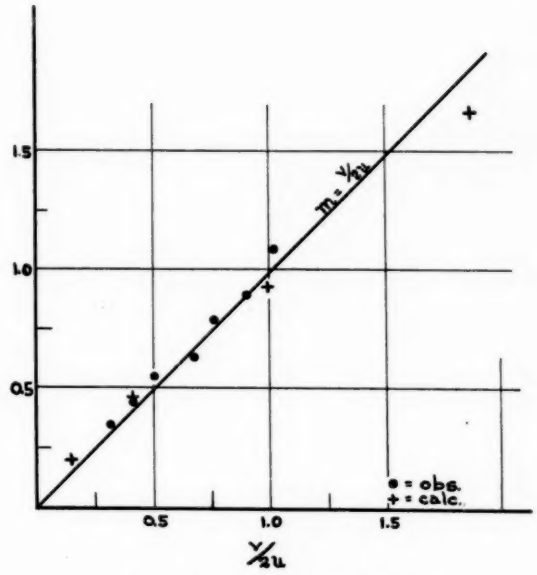


FIG. 7. Magnification (fixed focus).

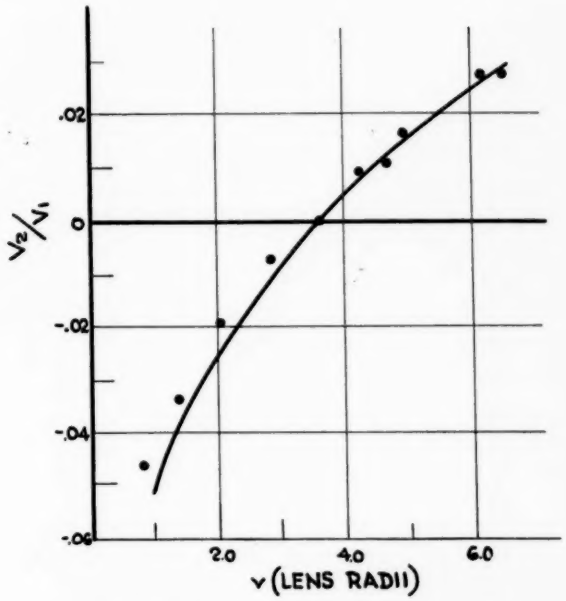


FIG. 8. Focusing relation (variable focus).

screen can conveniently be separated into two classes; axial aberrations and field aberrations. The former, chromatic and spherical aberration in the customary nomenclature, affect the quality of the image on the axis, as well as in other parts of the field, while the latter, in first line astigmatism and curvature of field, become noticeable only in the outer portions of the image. It will be seen that these constitute the most serious defects of the systems here studied.

An experimental investigation of all the image defects of electron optical systems has been carried out for a point source by Diels and Knoll,[4] to demonstrate their identity with the aberrations encountered in light optics. H. Johannson[5] has made a special study of the aberrations of the electron microscope. On the theoretical side, Glaser[6] has developed the theory of electron optical aberrations very completely, though not in a form in which it is directly applicable to our case, where the electrons enter the refracting field with infinitesimal velocities. Henneberg and Recknagel[7] have, making suitable simplifying assumptions, given a theoretical treatment of the chromatic (and spherical) aberrations of various types of image tubes.

It was not practical, in the present case, to carry out measurements of the axial aberrations, as these are not of sufficient magnitude to be readily determined except for so low accelerating voltages that no bright images could be obtained on the screen. For the measurement of the field aberrations, curvature and astigmatism, the

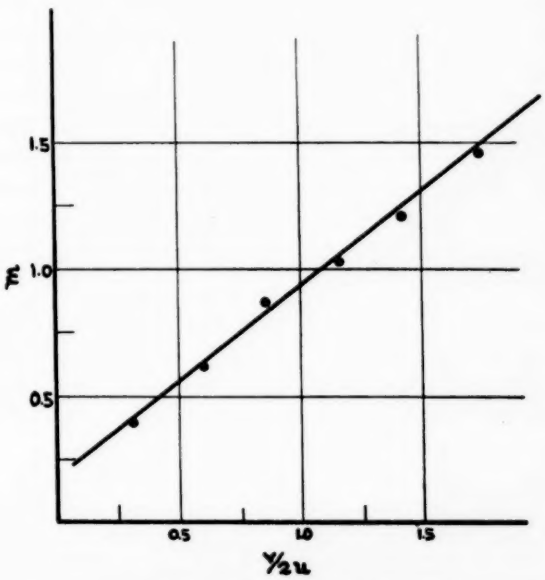
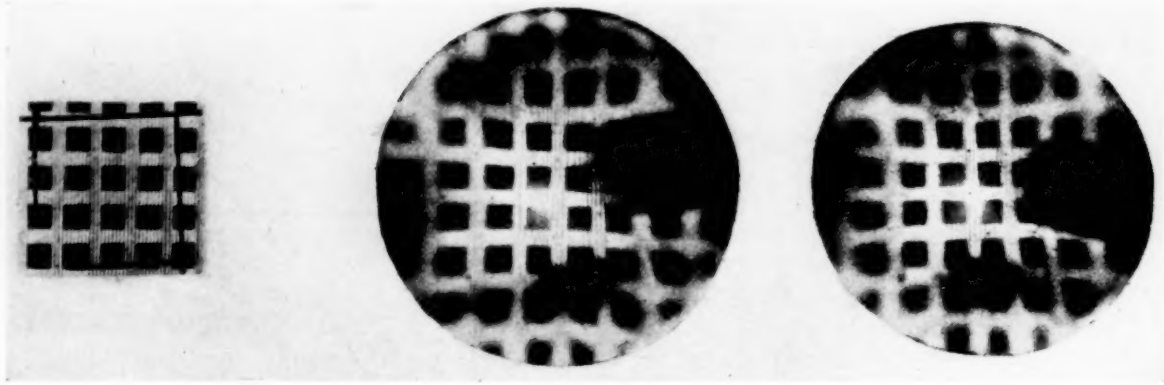


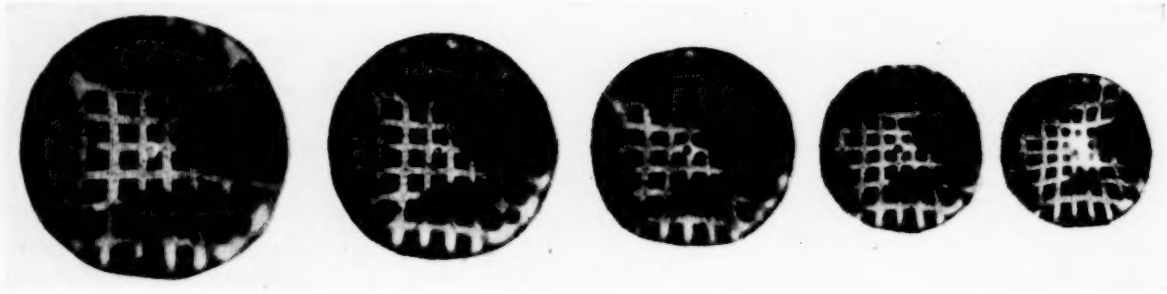
FIG. 9. Magnification (variable focus).

[4] K. Diels and M. Knoll, Zeits. f. Tech. Physik **16**, 617 (1935).
[5] H. Johannson, Ann. d. Physik **18**, 385 (1933); **21**, 274 (1934).
[6] W. Glaser, Zeits. f. Physik **80**, 451; **81**, 647; **83**, 104 (1933); **97**, 177 (1935); Ann. d. Physik **18**, 557 (1933).
[7] W. Henneberg and A. Recknagel, Zeits. f. Tech. Physik **16**, 230 (1935).



Object: $u=1.7R$ Image: $v=3.46R$ $R=1.5''$ $v=3.33\ R$

FIG. 10a.



$v=2.96\ R$ $v=2.51\ R$ $v=2.18\ R$ $v=1.80\ R$ $v=1.57\ R$

FIG. 10b.

FIG. 10. Sections through image space of flat cathode tube for various distances from lens v.

procedure was the following. In the tube of Fig. 2 the object distance was kept fixed at the value $u=1.8$ and pictures were taken of the image on the fluorescent screen at various distances from the lens; such a series of photographs is shown in Fig. 10. The object consisted of a grid of fine vertical lines superposed on an array of black squares. The point along a vertical line through the optical axis of the system at which the fine lines appear sharpest, indicates the location of a sagittal, that along a horizontal line through the axis where they appear sharpest, that of a tangential image point, just as in ordinary optics. Having located the sagittal and tangential image points in every plane, it is possible to construct also the sagittal and tangential image surfaces.

In order to calculate the axial aberrations, we consider pencils of rays leaving the center of the cathode with a fixed initial kinetic energy $e\Delta\Phi$ under different angles, e.g., 0°, 45° and 90°, and find their points of convergence on the axis.

To treat the first case, $\Theta=0$, we replace Φ in Eq. (13) by $\Phi+\Delta\Phi$ and then introduce as new dependent variable (which vanishes for $z=0$):

$$b=\frac{1}{2\left(\left[z+\frac{\Delta\Phi}{\Phi'(0)}\right]-\left[\frac{\Delta\Phi}{\Phi'(0)}\left[z+\frac{\Delta\Phi}{\Phi'(0)}\right]\right]^{\frac{1}{2}}\right)}-\frac{1}{r}\frac{dr}{dz}. \tag{16}$$

If b_0 is a solution of Eq. (15) corresponding to the case of zero initial velocity and $\Delta b=b-b_0$ is small up to the point where ray paths become substantially straight ($z=z_1$), we can write for $\Delta b(z_1)$:

$$\Delta b(z_1)=\int_0^{z_1} B(z')\exp\left(\int_{z'}^{z_1} A(z'')dz''\right)dz', \tag{17}$$

where $$A=-(1/z+\Phi/2\Phi)+2b_0, \tag{18}$$

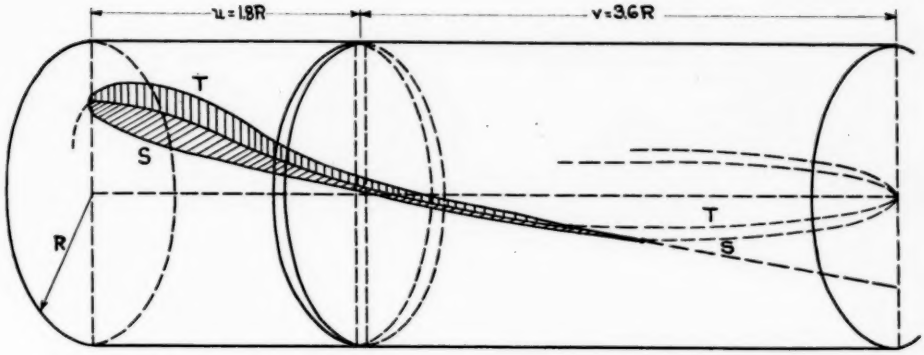


FIG. 11. Schematic diagram of lens structure and sagittal and tangential pencils.

$$B=\frac{\Delta\Phi}{\Phi}\left(b_0{}^2-\frac{db_0}{dz}-\frac{b_0}{z}-\frac{1}{4z^2}\right)-\frac{1}{4z^2}\left(\frac{\Delta\Phi}{\Phi'(0)z}\right)^{\frac{1}{2}}\times\left(1-\left(\frac{\Delta\Phi}{\Phi'(0)z}\right)^{\frac{1}{2}}\right)-\frac{b_0}{z}\left(\frac{\Delta\Phi}{\Phi'(0)z}\right)^{\frac{1}{2}}+\frac{\Phi'}{4z\Phi}\left(\frac{\Delta\Phi}{\Phi'(0)}\right)^{\frac{1}{2}}.$$

To find the solutions for rays with finite initial lateral velocity components, we introduce the variable $\Delta r=r-r_0$ where r_0 is the solution of the paraxial Eq. (13) for an equal initial lateral velocity and integrate the equation

$$\Delta r''=-\frac{\Phi''}{4(\Phi+\Delta\Phi)}\Delta r-\frac{\Phi'}{2(\Phi+\Delta\Phi)}\Delta r'-r_0\frac{\Phi''}{4(\Phi+\Delta\Phi)}\left[r_0{}^2\left(\frac{\Phi''}{4(\Phi+\Delta\Phi)}-\frac{\Phi_{IV}}{8\Phi''}\right)+2r_0{}'^2\right]-r_0'\frac{\Phi'}{2(\Phi+\Delta\Phi)}\left[r_0{}^2\left(\frac{\Phi''}{4(\Phi+\Delta\Phi)}-\frac{\Phi'''}{4\Phi'}\right)-r_0r_0'\frac{\Phi''}{2\Phi'}+r_0{}'^2\right]-r_0''\frac{\Delta\Phi}{\Phi+\Delta\Phi}. \quad (19)$$

This is obtained from Eq. (12a) if terms involving third powers of r and r' are retained. The proper initial conditions follow in every case from the form of the parabolic path described by the electron in the uniform accelerating field immediately in front of the cathode.

For the case of the fixed-focus tube with unity magnification and with $\Delta\Phi/V_1=0.0001$ (e.g., $\Delta\Phi=0.4$ volt, $V_1=4000$ volts), the axial deviations from the focal point corresponding to infinitesimal initial velocities are found to be 0.44, 0.24, and -0.03 lens radii for $\Theta=0°$, 45°, and 90°, respectively. As the solution of Eq. (13) for our particular case indicates that the vertex angle of a beam, with initial lateral velocity $[2(e/m)\Delta\Phi_r]^{\frac{1}{2}}$, is equal to 2.17 $(\Delta\Phi_r/V_1)^{\frac{1}{2}}$ the diameter Δ of the circle of least diffusion is found to be about 0.003 lens radii, or 0.004″(=0.1 mm), increasing approximately in proportion to the initial energy of the electrons. The formula given by Henneberg and Recknagel[7] yields $\Delta=2m\Delta\Phi/\Phi'(0)$ $=0.004$ lens radii or 0.006″(=0.15 mm). It is based on considering the image tube field as consisting of a uniform initial field terminated by a short lens. It is seen that the axial aberrations, while not important in the present case, may be of significance when the field before the cathode is reduced (as when a very small image is desired), or in cases when the initial velocities are on the average substantially larger and the accelerating voltage is reduced.

Of greatest importance are, however, the field aberrations, image curvature and astigmatism. To calculate these, the paths of a set of principal rays, corresponding to electrons starting from rest on the cathode at various distances from the axis, must first be obtained by integrating the equation

$$r''=-\frac{r\Phi''}{4\Phi}\left[1+r^2\left(\frac{\Phi''}{4\Phi}-\frac{\Phi_{IV}}{8\Phi''}\right)+2r'^2\right]-\frac{r'\Phi'}{2\Phi}\left[1+r^2\left(\frac{\Phi''}{4\Phi}-\frac{\Phi'''}{4\Phi'}\right)-r\frac{r'\Phi''}{2\Phi'}+r'^2\right], \quad (20)$$

which is found by substituting Eq. (9) in Eq. (12a), retaining terms up to the third power in r

and r'. Pencils of rays starting from the same point on the cathode, with infinitesimal lateral velocities, will in general come to a focus only if they lie either in the meridional plane or a plane normal thereto, the focal points being termed the tangential and sagittal image points, respectively. These image points, corresponding to all possible principal rays, form the tangential and sagittal image surfaces which are the object of the calculation. The separation of the two surfaces gives a measure of the astigmatism, their curvature determines the image curvature. Fig. 11 illustrates the tangential and sagittal pencils belonging to a particular principal ray, as well as the two image surfaces T and S.

To obtain the tangential image point corresponding to a given principal ray $r_0(z)$, which is a solution of Eq. (20) for $r_0'(0)=0$, we introduce the difference variable $\Delta r=r-r_0$ and, furthermore, the convergence variable

$$b=1/2z-(1/\Delta r)\Delta r', \tag{21}$$

which obeys the equation

$$\begin{aligned}\frac{db}{dz}=&-\frac{1}{4z^2}+\frac{\Phi''}{4\Phi}\left[1+3r_0^2\left(\frac{\Phi''}{4\Phi}-\frac{\Phi_{IV}}{8\Phi''}\right)\right.\\&\left.+2r_0r_0'\left(\frac{\Phi'}{2\Phi}-\frac{\Phi'''}{2\Phi''}\right)+r_0'^2\right]\\&+\frac{\Phi'}{2\Phi}\left(\frac{1}{2z}-b\right)\left[1+r_0^2\left(\frac{\Phi''}{4\Phi}-\frac{\Phi'''}{4\Phi'}\right)\right.\\&\left.+r_0r_0'\frac{\Phi''}{\Phi'}+3r_0'^2\right]-\frac{b}{z}+b^2.\end{aligned} \tag{22}$$

This can be integrated numerically in the same fashion as Eq. (15) giving the tangential image point on the principal ray. The calculation of the tangential image curvature on the axis can readily be reduced to quadrature if $\Delta b=b-b_0$, where b_0 is the solution of Eq. (15), is introduced in Eq. (22).

To find the sagittal image point, the second Euler equation, (12b), is utilized, considering the displacement y as infinitesimal; then Eq. (12b) becomes

$$\begin{aligned}y''=y'&\left[\frac{r_0'r_0''}{1+r_0'^2}-\frac{1}{2\varphi}\left(\frac{\partial\varphi}{\partial r}r_0'+\frac{\partial\varphi}{\partial z}\right)\right]\\&+\frac{(1+r_0'^2)}{2\varphi r_0}\frac{\partial\varphi}{\partial r}y.\end{aligned} \tag{23}$$

Introducing

$$b=1/2z-(1/y)y', \tag{24}$$

as well as the expansion of the potential (9), a convergence equation analogous to (22) is obtained:

$$\begin{aligned}\frac{db}{dz}=&-\frac{1}{4z^2}+\frac{\Phi''}{4\Phi}\left[1+r_0^2\left(\frac{\Phi''}{4\Phi}-\frac{\Phi_{IV}}{8\Phi''}\right)+r_0'^2\right]\\&-\frac{r_0'r_0''}{2z}+\frac{\Phi'}{2\Phi}\left(\frac{1}{2z}-b\right)\left[1+r_0^2\left(\frac{\Phi''}{4\Phi}-\frac{\Phi'''}{4\Phi'}\right)\right.\\&\left.-r_0r_0'\frac{\Phi''}{2\Phi'}\right]+br_0'r_0''-\frac{b}{z}+b^2.\end{aligned} \tag{25}$$

The numerical work and the obtention of the sagittal radius of curvature on the axis proceed as in the case of the tangential pencil.

In Fig. 12, the measured image points are compared with the calculated image surfaces,

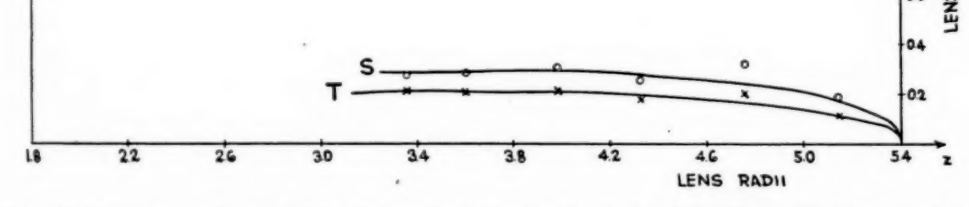


FIG. 12. Curvature of image field. S, calculated sagittal image surface; T, calculated tangential image surface; o, measured sagittal image points; ×, measured tangential points.

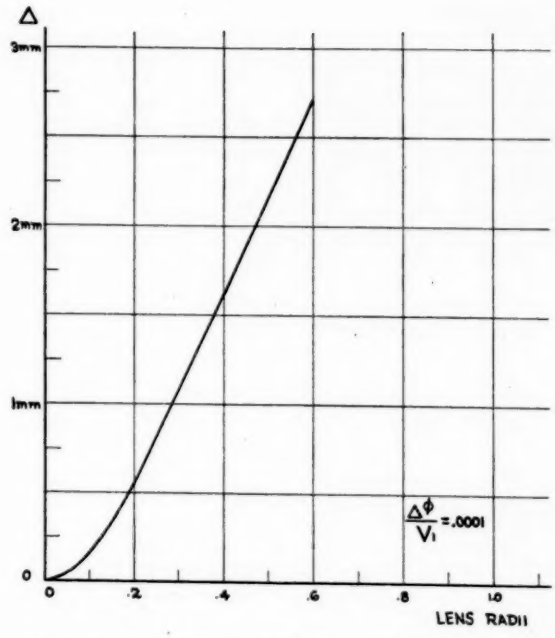
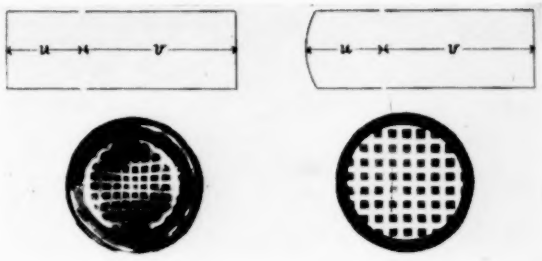


FIG. 13. Large diameter of ellipse of diffusion due to curvature of field as function of radial distance of object point.

computations being made for object points on the cathode separated by 0.2, 0.4 and 0.6 lens radii from the axis, as well as for the radii of curvature on the axis. Vertical and horizontal scales are identical so that the figure gives a true picture of the appearance of the image surfaces. Fig. 13 gives the corresponding large diameters of the ellipses of diffusion in the plane of the paraxial image point neglecting axial aberrations. The axial radii of curvature of the tangential and sagittal image surfaces are only 0.044 and 0.075 lens radii, or 1.7 and 2.9 mm, respectively.

It was found, as already reported in the earlier paper,[1] that this very serious defect could be largely removed by giving the cathode a curvature such that its center of curvature fell, approximately, into the center of the lens. The comparison of the images given by the flat and the curved cathode tube, shown in Fig. 14, brings this out strikingly. Without, as yet, giving a quantitative treatment of the curved cathode tube, the main reason for the improvement may readily be pointed out. By curving the cathode in the above mentioned fashion, the principal rays are practically undeflected in the first part of their path, and the sagittal and tangential pencils of all the rays encounter practically identical variations of potential. As, due to the low velocities of the electrons in this region, their paths are most easily influenced in the early portion of their transit, it is clear that this must minimize the astigmatism, i.e., the difference between the convergence of the tangential and sagittal rays, and similarly reduce the curvature due to the similar conditions along different principal rays.



Flat Cathode Curved Cathode

FIG. 14. Comparison of images from tubes with flat (left side) and curved (right side) cathodes.

In conclusion, the authors wish to express their thanks to Dr. V. K. Zworykin, the director of the Electronic Research Laboratory, and to the members of the laboratory staff whose cooperation made this work possible.

Directional Characteristics of Any Antenna Over a Plane Earth

W. W. HANSEN, *Stanford University, California*

(Received October 5, 1936)

Simple expressions for the field as a function of direction at distances of many wave-lengths from any antenna above a plane earth are obtained. Except near the surface of the earth the field varies as $1/r$. In agreement with results of others it is found that on the surface of the earth for values of kr, $(k=2\pi/\lambda)$, large compared to unity and yet small compared to the (complex) dielectric constant, the vertical component of the field varies as $1/r$ while if the distance is large compared to the dielectric constant the field decreases as $1/(kr)^2$. The various constants etc., depending on antenna form and ground characteristics, are shown to be connected with quantities previously used for the determination of radiation resistance so that the same numerical work used to find the total radiation from an antenna also determines its directional characteristics.

IN the present paper simple expressions valid at large distances are obtained for the intensity of radiation from any antenna over an arbitrary earth as a function of angle. This work is an extension of the theory reported in two recent papers.[1, 2] There exact formulae were given which enable the computation of the *total* power radiated from any antenna above a plane earth of arbitrary characteristics. Moreover it was shown that these formulae are eminently suitable for practical calculation. As an intermediate step in obtaining the total power radiated expressions for the fields were obtained but these expressions were integrals utterly impossible for numerical work; so little was said about them. We now find, however, that it is a simple matter to evaluate these expressions in terms of approximations valid at large distances and so we are able to get the directive patterns. Thus both the radiation resistance and the directive characteristics of any antenna are now obtainable from the same mathematical formalism and such numerical integrations etc., as are needed to give either of these properties also suffice to give the other.

To be more exact as to the results, we find three separate expressions for the fields at large distances. These expressions have different regions of validity and taken together cover most of the interesting cases. The first and leading term in the asymptotic expression gives the field many wave-lengths from the origin for all angles. At the surface of the earth, however, this expression vanishes and we need the next term of the series. Thus we are led to examine the propagation of the waves at the earth's surface. This has been done before by a number of people and the brief treatment here adds only a little that is new. Specifically, in addition to the expression mentioned above, we find two expansions good at the surface of the earth, one valid at a distance (in wave-lengths) from the antenna large compared to one and yet small compared to the square of the ratio of the wave-length in air to that in the ground; the other valid when the distance is large compared to both unity and this squared ratio.

To understand the proposed method of attack we need the previous results. It was shown that the electric field is given by

$$\mathbf{E}=-\frac{k^2}{2c}\sum_n\int_0^{i\infty+\pi/2}([f_{2,n}(\theta')+\alpha_2 f_{2n}(\pi-\theta')]\mathbf{A}_{2n}(\rho\phi z:\theta')+[f_{3,n}(\theta')+\alpha_3 f_{3,n}(\pi-\theta')]\mathbf{A}_{3n}(\rho\phi z:\theta'))\sin\theta' d\theta', \quad (1)$$

where the integration is from 0 to $\pi/2$ to $\pi/2+i\infty$ and the various symbols are defined as follows.

Let $\xi_n(\rho\phi z:\theta')=e^{i(kz\cos\theta'+n\phi)}J_n(k\rho\sin\theta')$, (2a)

and then $\mathbf{A}_{2n}=\nabla\times\mathbf{k}_z\xi_n=(1/k)\nabla\times\mathbf{A}_{3n}$,

$\mathbf{A}_{3n}=(1/k)\nabla\times\mathbf{A}_{2n}$, (2b)

[1] W. W. Hansen and J. G. Beckerley, Physics **7**, 220 (1936).

[2] W. W. Hansen and J. G. Beckerley, to appear in Dec. issue Proc. I.R.E.

where ρ, ϕ, z are cylindrical coordinates with origin at the surface of the earth and with the z axis vertical. The "oscillator strengths" f are given in terms of the current distribution $\mathbf{i}(\rho\phi z)$ as follows:

$$f_{2n}(\theta')=1/(2\pi k^2\sin^2\theta')\int \mathbf{i}\cdot\bar{\mathbf{A}}_{2n}d\tau,$$
$$f_{3n}(\theta')=1/(2\pi k^2\sin^2\theta')\int \mathbf{i}\cdot\bar{\mathbf{A}}_{3n}d\tau. \tag{3}$$

Finally the α are reflection coefficients which depend on the characteristics of the ground and on θ'.

$$\alpha_2=\frac{\mu\cos\theta'-(\epsilon'\mu-\sin^2\theta')^{\frac{1}{2}}}{\mu\cos\theta'+(\epsilon'\mu-\sin^2\theta')^{\frac{1}{2}}},$$
$$\alpha_3=\frac{\epsilon'\cos\theta'-(\epsilon'\mu-\sin^2\theta')^{\frac{1}{2}}}{\epsilon'\cos\theta'+(\epsilon'\mu-\sin^2\theta')^{\frac{1}{2}}}, \tag{4}$$

with
$$\epsilon'=\epsilon+(i\sigma/kc), \tag{4a}$$

where ϵ, μ, σ are the dielectric constant, permeability and conductivity, respectively, of the ground.

By varying θ' we get an infinite number of functions $\mathbf{A}_2$, $\mathbf{A}_3$, all of which are solutions of the wave equation. Thus Eq. (1) says that the vector potential is a sum of such solutions with the various solutions weighted by a factor g

$$g_{2n}(\theta')=f_{2n}(\theta')+\alpha_2(\theta')f_{2,\,n}(\pi-\theta'),$$
$$g_{3n}(\theta')=f_{3n}(\theta')+\alpha_3(\theta')f_{3,\,n}(\pi-\theta'). \tag{5}$$

We may note for future reference that since $\alpha(\pi/2)=-1$ we have $g(\pi/2)=0$.

Now the exact expression (1) for $\mathbf{E}$ is not usable because, for large values of r, the integrand oscillates rapidly and numerical integration is impossible. The problem, then, is to find an approximate expression for (1), valid for large distances from the antenna and not too hard to compute. We have found several methods for doing this. The one we will describe follows a scheme first introduced by Whittaker[3] and later used in connection with radiation problems by Weyl[4] and others. The wave functions $\xi_n(\rho\phi z:\theta')$

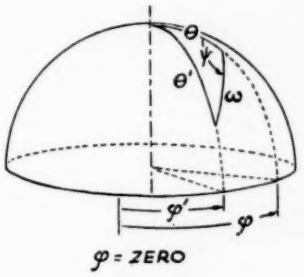


FIG. 1.

from which the $\mathbf{A}$ are generated are considered as made up of an infinity of plane waves going in all directions with properly chosen amplitudes and phases. Thus the field is expressed as a sum of plane waves. If now we go a long distance from the origin in a certain direction it turns out that the component plane waves going in, or nearly in, that direction have nearly the same phase and add up to be important, while the waves going in appreciably different directions have practically random phases and add to naught. Thus to find the field in a given direction we need only the amplitude of the component plane wave going in that direction and a simple expression for this amplitude is easily obtained.

To carry through the method sketched above we proceed as follows. First of all we avoid vector integrations by going back to the definition (2b) of the $\mathbf{A}$. Also we simplify the equations by considering for the moment only one term of (1).

$$\mathbf{E}=\frac{-k^2}{2c}\nabla\times\mathbf{k}_z\int_0^{i\infty+\pi/2} g_n(\theta')\xi_n(\rho\phi z:\theta')\sin\theta' d\theta'. \tag{6}$$

The wave function ξ may be built up by plane waves in the following manner:

$$\xi_n=\frac{1}{2\pi i^n}\int_{-\pi}^{+\pi} e^{ikX}e^{in\phi'}d\phi', \tag{7}$$

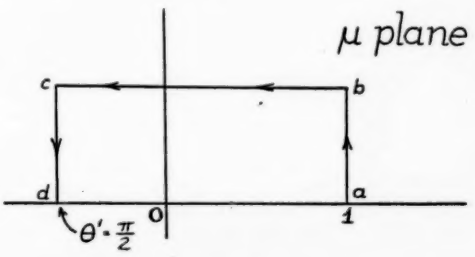


FIG. 2.

[3] E. T. Whittaker, Math. Ann. **57**, 333 (1902). See also, Whittaker and Watson, *Modern Analysis*, fourth edition (Cambridge, 1927), p. 388 and ff.

[4] H. Weyl, Ann. d. Physik **60**, 481 (1919).

where

$$X = z\cos\theta' + x\sin\theta'\cos\phi' + y\sin\theta'\sin\phi'. \quad (7a)$$

So we have

$$\mathbf{E} = \frac{-k^2}{2c}\frac{1}{2\pi i^n}\nabla\times\mathbf{k}_z\int_{-\pi}^{+\pi}\int_0^{i\infty+\pi/2} e^{ikX} \times g(\theta')e^{in\phi'}\sin\theta' d\theta' d\phi'. \quad (8)$$

Now if we wish the field $\mathbf{E}$ at a large distance from the origin in a direction specified by polar coordinates θ, ϕ it is convenient to state the direction θ', ϕ' in which one of the component plane waves is traveling by means of polar angles measured from an axis pointing in the θ, ϕ direction. We therefore introduce such angles, call them ω, ψ, defined by

$$\cos\omega = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\phi-\phi') \qquad \sin\theta'\sin(\phi-\phi') = \sin\omega\sin\psi \quad (9)$$

and illustrated in Fig. 1. Then on putting in expressions for x, y, z in terms of the spherical coordinates, r, θ, ϕ, with polar axis coincident with the z axis and making the substitution (9), (8) becomes

$$\mathbf{E} = \frac{k^2}{2c}\frac{1}{2\pi i^n}\nabla\times\mathbf{k}_z\int\int e^{ikr\mu}g(\theta')e^{in\phi'}d\mu d\psi, \quad (10)$$

where $\mu = \cos\omega$. As to the limits we first notice that, provided $z \neq 0$, it can be shown that in the original integral (1) the part of the path from $\pi/2$ to $\frac{\pi}{2} + i\infty$ gives a contribution to the total that vanishes like $r^{-5/2}$ and is therefore in general negligible when $kr \gg 1$. Neglecting this, then, the integration in (10) is over the half-sphere $0 < \theta' < \pi/2$, $-\pi < \phi' < \pi$.

We now perform the integration over μ after first distorting the path of integration as shown in Fig. 2. This distortion is allowable since $g(\theta')$ has no singularities inside the contour. The point of the distortion is that along the contour from a to b the integrand does not oscillate but decreases exponentially. The integration from b to c gives nothing because the integrand is vanishingly small and from c to d the integrand is zero because $g(\pi/2) = 0$. Moreover, if kr is sufficiently large the relevant region of integration is correspondingly close to $\mu = 1$ and $g(\theta')e^{in\phi'}$ may be considered as constant with value corresponding to $\omega = 0$, i.e., $\theta = \theta'$, $\phi = \phi'$, and then the actual integration is an elementary one, with result

$$\mathbf{E} \doteq \frac{k^2}{2c} i^{-n+1}\nabla\times\mathbf{k}_z\frac{e^{ikr}}{kr}g(\theta)e^{in\phi}. \quad (11)$$

Just how large kr must be is worth further consideration. The restrictions may be divided into three classes, those that depend on the method used, those that depend on the antenna form, and finally those that depend on the ground characteristics. For the first it is plain that we must have $1 \ll kr$ and it is also fairly evident that there must be a number of waves between $\omega = 0$ and $\theta' = \pi/2$; this leads to the restriction $1/(\pi/2-\theta)^2 \ll kr$. This last could perhaps be removed by a slightly changed derivation. Restrictions of the second and third types come from essentially the same requirement, namely that the average of $g(\theta')e^{in\phi'}$ over the region covered by the first wave on the ω, ψ sphere shall not be greatly different from the value at $\omega = 0$. Plainly this involves the ratio of the second derivative to the function itself. Considering the variation due to $e^{in\phi}$ it is readily verified that we must have $n^2 \ll kr$. Now in order to have an f_n of important size it is essential that the antenna extend at least far enough from the axis to have $k\rho \sim n$. In other words, measuring all distances in wave-lengths, we must be at a distance greater than the square of the distance from the z axis to the farthest part of the antenna. Much the same situation is met in considering variations in g that are due to changes in f. Thus it is impossible to get rapid variations of f without a high antenna and this leads to a requirement that kr be large compared to the square of the maximum value of kz on the antenna. These last two requirements and the very first one too are best lumped together into the single statement that kr must be large compared to $(kr_0+1)^2$ where r_0 is the radius of the smallest sphere centered on the origin that will enclose the antenna.[5] Finally we must consider

[5] This restriction on kr is very easily understood by considering the problem without ground and in spherical coordinates. There the field involves half-order Bessel functions, the order of the highest important one being

the changes of g due to the rapid decrease of $1+\alpha$ near $\theta'=\pi/2$. On computation it is found that this leads to $\epsilon'/(\epsilon'\mu-1)^{\frac{1}{2}}(\pi/2-\theta)\ll kr$ for the θ component and a similar expression for the ϕ component with ϵ' and μ interchanged. This last restriction keeps the forthcoming Eq. (12) from being valid at the surface of the ground and so keeps it from predicting zero field strength there.

The operation $\nabla\times$ is now performed and the deleted summation restored and we obtain

$$\mathbf{E}\cong\frac{+k^3}{2c}\frac{e^{ikr}}{kr}\sum_n(\mathbf{k}_\theta i^{-(n+1)}g_3(\theta)\sin\theta e^{in\phi}+\mathbf{k}_\phi i^{-n}g_2(\theta)\sin\theta e^{in\phi}). \quad (12)$$

Thus we have found an expression for the field strength as a function of angle, this expression being valid for large values of kr, i.e., many wave-lengths from the antenna. We may notice that if we compute the Poynting vector from (12) it is obviously in the radial direction and that on integrating over the half-sphere $0<\theta<\pi/2$ we get an expression for the energy radiated upward that is identical with Eq. (17) of reference 2.

As remarked above Eq. (12) is invalid at $z=0$ and as this is often the most interesting region we next look for an expression good there.

The problem of the variation with kr of the field at $z=0$ has already been treated at length by many writers in the case where the transmitter is a dipole. In fact there exist fairly simple approximate formulae from which fields at all distances can be obtained. It would be in the highest degree surprising if the field due to a complex antenna varied in an appreciably different way at distances $kr\gg(kr_0)^2$, though just what constant factor should be used is perhaps not quite plain. The analysis to be presented verifies this guess and also gives an expression for the constant to be used when the antenna is not an infinitesimal dipole. In one way, however, we have not gone as far as possible for though a single formula valid for all kr probably exists we find instead two approximate formulas, good in different regions of kr.

Before looking for general expressions, we may glance at three rather obvious special cases. First if $\epsilon'=\mu=1$, i.e., no ground at all, then $\alpha_2=\alpha_3=0$, the g are not in general zero at $\theta'=\pi/2$ and (12) is both valid and useful, though the validity is not obvious without a slightly changed derivation. Second, let $\mu=1$ and $|\epsilon'|\to\infty$ for example, by letting the ground become a perfect conductor. Then, as inspection of (4) will show, $\alpha_3=-1$, $\alpha_3=+1$. This makes $g_2(\pi/2)=0$, $g_3(\pi/2)=2f_3(\pi/2)$ so causing (12) to indicate that the ϕ component of $\mathbf{E}$ vanishes at the earth's surface and that the θ component is just twice what it would be without ground. Both of these conclusions are obviously correct, the first because there can be no tangential field at a perfect conductor, the second by the elementary theory of images. Finally, the third limiting case of $\epsilon'=1$ $\mu\to\infty$ gives the same result as the second but with the θ and ϕ components interchanged.

We will now proceed with the problem of finding expressions for the field at $z=0$. We will carry out the work in detail only for $\mathbf{A}_2$, merely stating the result for $\mathbf{A}_3$. First of all the expression (1) is, strictly speaking, valid only for z greater than the highest part of the antenna. For $z=0$ it should be replaced by a slightly different equation derived from Eq. (8a), reference 1. This differs from (1) in that, for $z=0$, there occurs instead of $f(\theta')+\alpha f(\pi-\theta')$ the factor $f(\pi-\theta')(1+\alpha)$. Now it is convenient to make the change of variables $L=\sin\theta'$, $M=\cos\theta'$. On doing this the factor $\sin\theta' d\theta'$ becomes LdL/M and we have

$$\mathbf{E}\cong\mathbf{k}_\phi\frac{k^3}{2c}\int_0^\infty J_n'(krL)f_{2n}(\theta')(1+\alpha_2)\frac{L}{M}dL, \quad (13)$$

where $\mathbf{A}_2$ has been replaced by an approximate expression valid for large values of $k\rho L$. Also ρ has been replaced by r as the two are identical for $z=0$. Next we perform the addition $1+\alpha_2$ and find

$$1+\alpha_2=2M\frac{1}{(1-L^2)^{\frac{1}{2}}+(1/\mu)(\epsilon'\mu-L^2)^{\frac{1}{2}}}. \quad (14)$$

about $l=kr_0$. Then provided $kr\gg l^2$ or $kr\gg(kr_0)^2$ the Bessel functions are well approximated by single terms of Hankel type asymptotic expansions and these combine to give an r factor of the form e^{ikr}/r. But for smaller values of kr this is not possible; from a mathematical point of view because the asymptotic expansion is not accurate, from a physical point of view because the waves really are longer near the antenna.

Thus $1+\alpha_2$ goes to zero like M at $L=1$ so canceling the $1/M$ in (13) and avoiding an infinity in the integrand. This leads us to define the function[6]

$$h_2=f_2(1+\alpha_2)/M, \tag{15}$$

with which we can rewrite (13) as

$$\mathbf{E}=\mathbf{k}_\phi\frac{k^3}{2c}\int_0^\infty h(L)J_n'(krL)LdL. \tag{16}$$

To get a qualitative understanding of the next step we may profitably consider the integral as equal to the sum of an alternating series, terms of which are obtained by integrating between successive nodes of the integrand. Now for large kr the Bessel function oscillates rapidly so the terms are very nearly equal in absolute value and cancel to a high order except in those regions where h is changing rapidly. This argument leads us to examine the behavior of h with the object of finding where it is changing rapidly and then to transform the integral in such a way as to make use of the above observation.

Examining (15) and (14) we see that h has a sharp peak at $L=1$ and also a rise when L is near $(\epsilon'\mu)^{\frac{1}{2}}$. This latter we will henceforth neglect because, for a number of reasons, it is unimportant when $\epsilon'\mu$ is large compared to one, and we will confine ourselves to this case. Considering the region near $L=1$ we see that h_2 reaches a peak value of $2f_2(\pi/2)\mu/(\epsilon'\mu-1)^{\frac{1}{2}}$ at $L=1$ and drops to half this value when $\delta=1-L=(\epsilon'\mu-1)/\mu^2$. When δ is less than this value and also less than one, we find by the binomial expansion that h_2 is well approximated by

$$h_2\leqq 2f_2(\pi/2)\frac{\mu}{(\epsilon'\mu-1)^{\frac{1}{2}}}\left[1-\frac{(2\mu)^{\frac{1}{2}}}{(\epsilon'\mu-1)^{\frac{1}{2}}}\delta^{\frac{1}{2}}\cdots\right]. \tag{17}$$

On the other hand if μ is large so that δ can be small compared to one and yet too large for (17) to hold, then

$$h_2\leqq(2)^{\frac{1}{2}}f_2(\pi/2)\frac{1}{(\delta)^{\frac{1}{2}}},\quad \frac{\epsilon'\mu-1}{\mu^2}\ll\delta\ll 1. \tag{18}$$

We can now evaluate (16). If it happens that $1\ll kr\ll\mu^2/(\epsilon'\mu-1)$ then there will be only a fraction of an oscillation of the integrand in the region covered by (17) and a considerable number in that where (18) is valid. We therefore put (18) in (16) and, since the important region is near $L=1$, we are also allowed to approximate the Bessel function by its asymptotic expansion, provided $kr\gg n^2$. Doing this and using δ as an integration variable, we are led to integrals of the general form

$$\int F(\delta)\sin kr\delta\frac{d\delta}{(\delta)^{\frac{1}{2}}}, \tag{19}$$

where F is a slowly varying function of δ. These are easily evaluated by using a theorem due to Bromwich,[7] which says that if $-1<\beta<1$ and certain other conditions are fulfilled, then when $\nu\to\infty$

$$\int_0^\infty x^{\beta-1}F(x)\sin\nu x dx=\nu^{-\beta}F(0)\Gamma(\beta)\sin\tfrac{1}{2}\beta\pi. \tag{20}$$

We see that our qualitative prediction that the value and manner of variation of h near $L=1$ would be critical is verified. The actual evaluation is now an elementary matter and we easily find after restoring the deleted summation and the terms in $\mathbf{A}_3$,

$$\mathbf{E}\leqq\frac{k^3}{2c}\frac{e^{ikr}}{kr}2\sum e^{in\phi}(\mathbf{k}_\theta f_{3n}(\pi/2)i^{-(n+1)}+\mathbf{k}_\phi f_{2n}(\pi/2)i^{-n}). \tag{21}$$

The restrictions on kr for the two terms are different and in fact mutually contradictory. If $(n+1)^2\ll kr\ll\epsilon'^2$ the first term is correct, but it will be impossible to also have $kr\ll\mu^2$ (and also $\epsilon'/\mu\gg 1$); so the second term is not to be used. In practice it seems likely that the first term only will be of use as high values of μ do not occur naturally.

We may compare the result of (21) with the various special cases mentioned above. If $\epsilon'\to\infty$ the first term of (21) is valid for all values of kr and gives a field in the θ direction just twice what would exist without ground. If $\mu=1$, $\epsilon'\to 1$ then $\epsilon'^2/(\mu\epsilon'-1)\to\infty$ but in spite of this

[6] Our h_2 is equivalent to Sommerfeld's $1/N'$. Moreover for a magnetic dipole at $z=0$, $f_2=$constant, in which special case (16) will be recognized as identical with expressions for essentially the same thing obtained by Sommerfeld. See, for example, Riemann-Weber, Vol. II, p. 565. There is a similar connection in the electric dipole case.

[7] T. J. Ia' Bromwich, *Theory of Infinite Series*, 174.

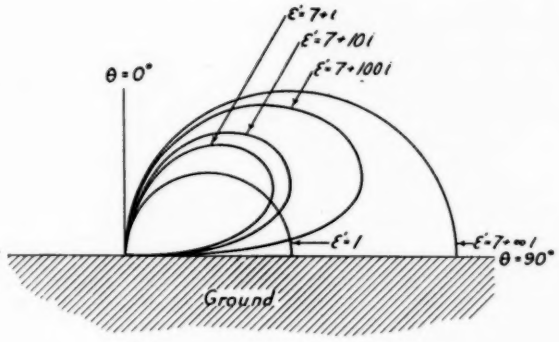


FIG. 3.

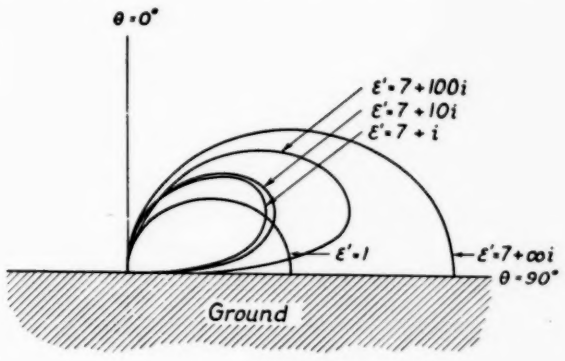


FIG. 4.

(21) is not expected to be valid because of the neglect of the peak in h at $L=(\epsilon'\mu)^{\frac{1}{2}}$. It is easily verified, however, that in the limit $\epsilon'\to 1$ this neglect should make the result too high by a factor two as in fact it does.

If in addition to being large compared to $(n+1)^2$, kr is also large compared to $\mu^2/(\epsilon'\mu-1)$, then approximation (17) is the proper one to use. The first term of this gives nothing but the second term after manipulations similar to those described above gives

$$\mathbf{E}\cong\frac{k^3}{2c}\frac{e^{ikr}}{(kr)^2}\frac{2}{\epsilon'\mu-1}\sum_n e^{in\phi}(\mathbf{k}_\theta f_{3n}(\pi/2)\epsilon'^2 i^{-(n+1)} + \mathbf{k}_\phi f_{2n}(\pi/2)\mu^2 i^{-n}), \quad (22)$$

the first term of which is valid for $(n+1)^2\ll kr$ and $\epsilon'^2/\epsilon'\mu-1\ll kr$; for the second term ϵ' in these conditions is replaced by μ.

Taken together the expressions (21) and (22) describe the behavior of $\mathbf{E}$ at $z=0$. In agreement with previously known results they say that, if $\mu=1$ and $\epsilon'\gg 1$, the field is almost entirely perpendicular to the earth's surface, that it varies like $1/kr$ for $kr\ll\epsilon'$ and like $1/(kr)^2$ for $kr\gg\epsilon'$. Moreover, in the latter case, the field is about ϵ'/kr times the strength it would have if the inverse distance formula (21) held so that at any fixed distance increasing ϵ' will increase the field in proportion until finally (22) is supplanted by (21). In the limiting case $\epsilon'\to\infty$ the perpendicular field is just twice what it would be without ground (first term of (21)) and the tangential component is zero (second term (22)). Finally we may note that the restriction $(\pi/2-\theta)\gg\epsilon/((\epsilon'\mu-1)^{\frac{1}{2}}kr)$ prevents (12) from ever predicting a lower intensity near $z=0$ than (21) and (22) give there. In fact it always gives a higher intensity, thus the field strength increases as one goes from the earth's surface upward.

The one point that is perhaps new and useful is the proof that, as one might expect, the intensity at $z=0$ is proportional to $f(\pi/2)$, i.e., to the intensity that would exist if the ground were removed. Thus to find the intensity at $z=0$ due to a given antenna it is only necessary to multiply the current distribution by $\mathbf{A}(\rho\phi z:\pi/2)$ and integrate. This is especially easy because when $\theta'=\pi/2$ all the $\mathbf{A}$ are independent of z and this greatly simplifies calculations.

To illustrate the use of (12) we have prepared Figs. 3 and 4 which give, respectively, the field as a function of angle for a vertical dipole at the surface of the earth[8] and a vertical quarter-wave antenna grounded at the lower end. The needed f values we have taken from work done in connection with reference 2. We have material from the same source for a number of other directive patterns, but these do not have symmetry about the z axis and so are rather harder to represent graphically. For this reason they are omitted.

[8] An expression from which Fig. 3 could be obtained was found by Weyl, see reference 4. Further work has been done later, see especially M. J. O. Strutt, Ann. d. Physik **1**, 721 (1929) and **4**, 1 (1930).

The Thermal Response of the Skin to Radiation*

JAMES D. HARDY AND THEODORE W. OPPEL, *Cornell University Medical College, New York City*

(Received October 19, 1936)

The thermal sensitivity of skin to radiation has been studied with a new technique and the temperature changes in the skin surface have been accurately measured. The change in skin response of both whites and negroes to visible, infrared radiation, and sunlight has been measured. Results show that the sensation evoked by the radiations depends upon the optical properties of the skin. White subjects were found to be more sensitive to heat than negroes. The effect of the exposed area upon the sensitivity was observed. The postulate that the magnitude of the sensation depends upon two factors, nerve impulse frequency and central association, was found to fit the present data. A formula for calculating the degree of association is proposed. The results show that the sensitivity to radiation increases as the exposed area is increased up to the point where the stimulus has dropped to 0.0004 gram cal./sec.·cm². A radiation weaker than this will evoke no sensation regardless of the size of the area exposed. This was taken to be the threshold of the excitation of an end organ. The temperature change in the skin surface with this radiation rate was measured to be a rise in temperature at the rate of 0.0008°C per second; sensation is perceived after 3 seconds of exposure. The relationship of the temperature change of the skin was studied. It was found that neither the *change* (ΔT) nor *rate of change* ($\Delta T/\Delta t$) of skin temperature is adequate to explain the sensory effect. An hypothesis is proposed which attributes the sensation to a differential change in the temperature of the blood vessel net work within 0.1 mm of the skin surface and another at 0.3 mm.

MODERN calorimetric studies have shown that the energy exchanges of the human body obey the first law of thermodynamics. Since the experiments of Atwater, Rosa and Benedict,[1] the maintenance of a balance between the heat produced within the body and that eliminated from the body surface has been recognized as one of the most important conditions for normal physiological function in man. The constancy of the level of body temperature (average 37°C) has long been used as an index of health although careful measurements[2] have revealed a diurnal variation of the internal body temperature of about 1°C, the temperature being about 36.5°C in the morning hours and increasing to a high level of 37.5°C in the evening. A resting temperature of more than half a degree above the normal average is, however, usually indicative of a disturbance of a pathological nature.

Body heat production arises as the result of the oxidation of food and of the exothermic chemical reactions in the body having to do with muscular work and other metabolic processes. The heat from the various organs is distributed over the body by the blood stream and finally brought to the skin surface from which it is eliminated. The rate of heat production is known to vary with emotion,[3] exercise, state of health, age, size, sex, the amount and kind of food in the stomach, temperature, etc. A comparison of the metabolic rates is generally made under conditions defined as "basal," that is with the subject awake and resting, 18 hours after the last meal. The heat production under these circumstances is known as the "basal metabolic rate" and amounts to about 38 kilogram calories per hour per square meter of body surface for a normal man and about 6 percent less for a woman. During severe exercise the heat production may *rise* as high as ten times the basal rate but in normal men the heat production cannot be *lowered* below the basal level even though the body may be receiving from external sources more than sufficient energy to maintain body temperature and the vital processes.

It is now generally recognized that the body is provided with a nervous mechanism for the control of body temperature.[4] The location of the thermostatic mechanism is usually given as the lower portion of the third ventricle of the brain.

* From the Russell Sage Institute of Pathology in Affiliation with the New York Hospital, and the Department of Medicine, Cornell University Medical College, New York.

[1] Atwater and Benedict, Carnegie Institution of Washington, Publication No. 42 (1905); Atwater and Rosa, Report of Storrs Agricultural Experimental Station (1907), p. 212.

[2] T. Deighton, "Physical Factors in Body Temperature Maintenance and Heat Elimination," Physiol. Rev. **13**, 431 (1933).

[3] E. F. Du Bois, *Basal Metabolism in Health and Disease* (Lea and Febiger, 1936).

[4] C. J. Martin, "Thermal Adjustment of Man and Animals to External Conditions," Lancet **2**, 561 (1930).

This brain area responds to thermal changes in the skin by means of special receptors for temperature, and to temperature changes in the blood passing into the brain. It is the function of this mechanism to so regulate heat production and loss that a uniform body temperature of 37°C is maintained. The control of heat production is limited to increasing the energy produced. When the body becomes cold, spontaneous muscular contraction occurs in the form of shivering and teeth chattering, and the heat production may rise several times the basal value. The first response of man on exposure to a cold environment is the diversion of the blood stream away from the skin by constricting the peripheral blood vessels. This affords the body a considerable layer of insulating matter (the thermal conductivity of flesh is the same as that of cork) and in effect "converts the integument into a suit of clothes." Continued exposure, however, will so cool the blood flowing to the brain that shivering is brought on. If by means of an artificial injection of immune typhoid vaccine or the natural infusion of malarial parasites in disease, the thermostatic mechanism be "set" on a higher temperature plane, all of the above reactions to cold are brought on even at normal body temperature. Thus one has a "chill" necessarily followed by a fever.

Upon exposure to a warm environment the skin blood vessels are dilated to facilitate the passage of heat to the surface. Continued exposure activates the sweat glands so that body heat is entirely dissipated by evaporation of moisture when the external temperature reaches 36°C or 95°F. In this manner, under all circumstances of internal heat production and usual external temperature change, the body temperature is maintained within normal limits. The present investigation is concerned with the sensitivity of the part of the regulatory mechanism which responds to heat and thermal radiation.

Du Bois, Hardy, and their co-workers[5] have shown that under normal conditions, 24 percent of the heat lost from the body is due to the evaporation of moisture from the lungs and skin, between 55 percent and 65 percent is due to the infrared radiation from the surface, and the remaining 15 to 20 percent to convection and conduction. It is evident that the major portion of the heat loss is dependent upon the level of skin temperature, a fact which would lead one to conclude that, even within the comfort zone (23°C to 27°C), a systematic change in the skin temperature with environmental temperature is brought about by means of a subconscious mechanism. Such a dependence of skin temperature upon environmental change is well known and the present interest in this phenomenon is the manner in which this regulation occurs.

The radiation to which man is exposed in his natural environment comprises a spectral range from 0.29μ in the ultraviolet to 20μ in the infrared. Physiologically, this range may be divided into four regions, namely the ultraviolet ($\lambda < 0.36\mu$), the visible ($0.36\mu < \lambda < 0.8\mu$), the penetrating infrared ($0.8\mu < \lambda < 3\mu$), and the nonpenetrating infrared ($\lambda > 3\mu$). The infrared spectrum is so divided because of the intense absorption bands of water and organic compounds at about (3μ) and beyond. The action of ultraviolet light upon the skin is largely of a photochemical character and this type of radiation does not occur naturally in sufficient amounts to affect the thermal mechanism of the body. The effects of the longer waved radiation is, insofar as is now known, entirely thermal and the spectral divisions are made on the basis of the difference in the optical properties of the skin for each region. The sun, which is the most important radiator in man's environment, has an energy distribution at the earth's surface on a clear day about as follows: 5 percent ultraviolet, 40 percent visible light, 54 percent penetrating infrared and 1 percent nonpenetrating infrared. The usual hot bodies in man's surroundings radiate most of their energy in the near infrared, and all objects including man himself radiate in the nonpenetrating infrared. The importance of these radiation exchanges is realized from a consideration of the fact that while lying under the tropical sun, a white man absorbs about three times as much heat from the sun as his basal heat production, and when exposed to zero weather loses heat at about twice his basal rate by radiation alone. This wide range of external conditions and the narrow limits of normal body temperature make severe demands upon a ther-

[5] E. F. Du Bois and J. D. Hardy, Proc. Nat. Acad. Sci., Washington (1935).

mostatic mechanism which is adequate for the control of the loss and production of body heat. The excellence of such a mechanism, in common with all thermostats, depends upon the rapidity of its response and the sensitivity of its detector of thermal changes. The present study concerns itself with the quantitative determination of the sensitivity of the human skin to thermal changes produced by radiation, and with a correlation of these data with the subjective sensation of warmth.

Many experiments have been performed to determine the sensitivity of the skin to heat and cold.[6] The methods usually consist of placing hot and cold objects upon the skin and following the external and internal thermal changes by means of thermocouples buried in the tissue at different depths. From such experiments the nerve endings responsible for temperature sense have been identified as the Krause and Ruffini organs located 0.1 mm and 0.3 mm, respectively, below the skin surface. A satisfactory idea of the thermal changes responsible for sensation can hardly be obtained from experiments of this type as the proximity of the sensitive endings to the skin surface does not permit the measurement of the changes in thermal gradients by means of wire thermocouples. Furthermore, the presence of foreign material near the end organs is known to inactivate them. Also, the use of the thermocouples for the measurement of surface temperature is open to criticism.[7] It is not surprising, therefore, that few definite conclusions could be drawn from data obtained with a thermocouple technique. The careful work of Bazett and his co-workers, using thermocouples, demonstrates the inadequacy of the method. Other experiments have been made with radiant energy to determine the maximum intensity which the skin can bear.[8] The function of skin pigment in relation to its effect on the absorbed radiation has received scant attention and it is obvious that such studies must be made before the question of skin sensitivity to radiation can be approached.

The thermal effect of incident radiation depends upon the amount of energy absorbed by the skin surface and the amount which penetrates into the body beyond the depth of the sensitive endings. The question of the reflecting power of the skin and its opacity to penetrating infrared and nonpenetrating infrared has been studied and it has been found that white and negro skin have the same optical properties in these regions.[9] That is, human skin absorbs 95 percent of all radiation of wave-length greater than 3μ within 0.1 mm of the skin surface and reflects less than 1 percent; between 1μ and 3μ the reflecting power of white and negro skin rises to about 18 percent and of the absorbed energy only 1 percent reaches a depth of greater than 3 mm. For the visible spectrum the reflection and penetration of white and negro skin have received less study. It is known that the reflecting power of white skin for visible light is approximately 35 percent, and that the penetrating power of visible radiation is greater than that for any other radiation in the range under discussion.[10]

Experimental

I. Dependence of sensation upon wave-length of stimulating radiation

The sensitivity of the skin was investigated for sunlight, visible, penetrating infrared, and nonpenetrating infrared radiation. The sources of radiation were obtained in the following manner. The sunlight was directed into the laboratory by means of a heliostat. The radiation was reflected from two first surface silvered mirrors, and its intensity was controlled with a diaphragm which varied the effective aperture. The visible light was obtained from a 1000-watt bulb after other radiations had been largely filtered out by passing the light through 9 cm of an 0.8 percent aqueous solution of copper sulphate. The spectral energy distributions of the artificial sources were determined with a spectrometer, and it was found that the visible radiation contained a band of near infrared between 1.0μ and 1.2μ amounting to 24 percent of the transmitted light, as shown in Fig. 1. Attempts to increase the relative

[6] H. C. Bazett, B. McGlone, R. G. Williams and H. M. Lufkin, Arch. Neurol and Psychiat. **27**, 489 (1932).

[7] J. D. Hardy, "Radiation of Heat from the Human Body II," J. Clin. Invest. **13**, 605 (1934).

[8] C. Sonne, "The Mode of Action of the Universal Light Bath V," Acta. Medica. Scand. **54**, 374 (1921).

[9] J. D. Hardy and C. Muschenheim, "The Radiation of Heat from the Human Body IV and V," J. Clin. Invest. **13**, 817 (1934) and **15**, 1 (1934).

[10] H. Laurens, *Physiological Effect of Radiant Energy*, Chap. II. (Chemical Catalog Co., New York, 1933).

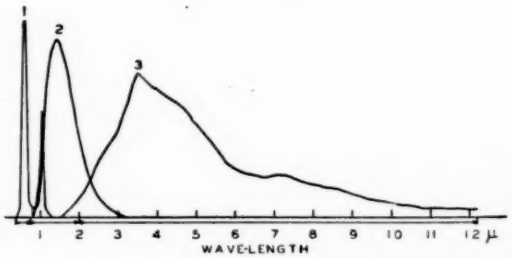


FIG. 1. Spectral energy distribution of sources of radiation used. Curve 1, visible light; curve 2, near or penetrating infrared; curve 3, far or nonpenetrating infrared.

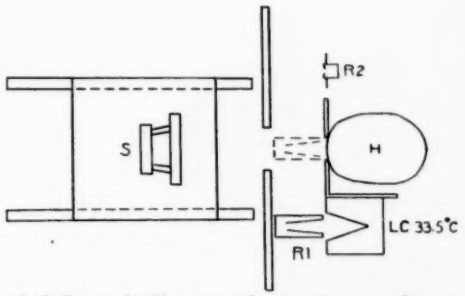


FIG. 2. Schematic diagram of apparatus used to measure minimum detectable energy and elevation of skin temperature during irradiation. *S*, source of radiation; *H*, subject's head; R_2 thermopile for measuring intensity of incident radiation; R_1 thermopile for measuring skin temperature; LC, Leslie Cube.

proportion of visible radiation resulted in greatly decreasing the intensity of visible red rays. A 1000-watt lamp also furnished the near infrared radiation after the rays had been passed through Corning's "heat transmitting" filter. The far infrared energy was obtained from a 250-watt electric stove, 12 cm in diameter, which was heated to a temperature between 300°C and 400°C before use.

The reflecting power of white and negro skin was measured for these radiations by comparing the diffuse reflection from the skin with that from a scraped block of magnesium carbonate. A thermopile was placed at a distance of 7 cm normal to the skin surface and the skin illuminated with very weak radiation from an angle of about 60°. The diffusely reflected light from the skin surface was compared with that from the carbonate which has been found to scatter in a manner similar to that of the skin. The results are recorded in Table I.

These values are in good agreement with the work of other authors.[10]

In Fig. 2 is shown a diagram of the apparatus used to determine the minimum quantity of energy necessary to stimulate a sensation. A convenient sized aperture, 14.5 cm², in the small screen limited a definite portion of the skin surface to the radiation. This aperture was so arranged that the test subject's head would fit comfortably behind the screen when the subject was seated. The large screen separated the subject from the operator and protected the measuring instruments from the temperature effects of the radiators. The source of radiation *S*, was mounted on a roll table so that the intensity of the stimulus could be altered by moving the source toward or away from the subject. A shutter which was operated silently was interposed between the subject and the radiator. Care was taken at all times that the subject should not know, by other means than heat perception, when stimulation was taking place. The advantage of this method over those in which the subject himself was the judge of his own perception is obvious.

TABLE I. *Diffuse reflection from skin.*

Surface	Visible Radiation	Sunlight	Penetrating Infrared	Non-penetrating Infrared
$MgCO_3$	86%	81%	73%	4%
White Skin	35	30	20	1
Negro Skin	16	17	17	1

Tests for the minimum perceptible amount of radiation were made in the following manner. The subject seated himself behind the screen with his head in position for some minutes in order to facilitate concentration. The operator then removed the shutter and irradiated the exposed skin surface with a stimulating amount of radiation for 3 seconds. This time of irradiation was chosen because it was found that if the stimulus could evoke a sensation it would do so within this period. For strong stimuli the subject signaled his perception within a second of the opening of the shutter, but for the faint sensations near the threshold the signal was given within 2 to 4 seconds after the beginning of the exposure. By placing this restriction on the time of the response false signals were largely eliminated. The intensity of the stimulus was decreased by steps and the test repeated until the smallest intensity was determined to which the subject responded each time the shutter was opened. The

TABLE II. *Minimum stimulus in g·cal./sec.·cm².*

Subject	Sunlight*			Visible Light			Penetrating Infrared			Nonpenetrating Infrared		
Whites	I	II	III	I	II	III	I	II	III	I	II	III
Trial 1	0.0063	0.0075	0.0063	0.0082	0.0082	0.0088	0.0045	0.0043	0.0038	0.0024	0.0034	0.0015
2	.0085	.0083	.0080	.0080	.0086	.0092	.0045	.0045	.0039	.0026	.0029	.0016
3	.0080	.0085	.0091	.0081			.0049	.0048		.0030	.0030	
4				.0088				.0049				
Average		.0081			0.0086			0.0045			.0025	
Energy Absorbed		.0056			.0056			.0037			.0025	
Negroes	I		II	I		II	I		II	I		II
Trial 1	0.0077		0.0074	0.0039		0.0048	0.0061		0.0064	0.0028		0.0019
2	.0066		.0078	.0054		.0044	.0062		.0064	.0040		.0020
3	.0060		.0092	.0035		.0039	.0063			.0039		
4	.0084			.0042			.0064			.0039		
Average	.0072		.0081	0.0043			0.0063			0.0031		
Energy Absorbed	.0060		.0065	.0037			.0049			.0031		

* Area irradiated = 23.8 cm².

factors of concentration and fatigue were found to be of considerable importance so that a large number of measurements could not be made on the same subject at the same time. After these factors were taken into consideration two or three complete tests were sufficient to determine the minimum detectable energy.

As soon as the operator had located the position of the source for the minimum sensation the subject removed his head and the radiometer R_2, was placed in the aperture. R_2 was provided with a metal flange which rested against the surface of the small screen in such a way that the thermopile had its absorbing surface in the same plane as that the skin had occupied. The sensitive element of R_2 was an eight junction circular thermopile (Coblenz type) whose absorbing surface was 9 mm in diameter and was blackened with India ink. The calibration of this instrument was carried out by means of a Leslie Cube. The voltage generated by the instrument was measured with a Leeds and Northrup potentiometer (Type K_2), the sensitivity of the device being such that 0.000126 g·cal./sec.·cm² produced 0.00001 volt.

(The calibration of this thermopile as recorded is in error by a factor 3.2. Thus, 0.000039 gram cal./sec./cm² produced 0.00001 volt, is the correct calibration. Although the energy values in the article are too great by this factor, interpretations of tables and curves are not changed. The one energy value whose absolute magnitude is of importance is that for the minimum detectable radiation which the body can perceive, which is 0.00012 gram cal. per second per cm² instead of 0.0004.—Authors.)

Tests were made on twelve subjects who showed moderate individual differences, but the ratios of the values for the different radiation bands were the same for all. The sensitivity values for three white subjects and two colored subjects shown in Table II are representative. It will be seen from the table that for the white subjects the important difference is a gradual increase in sensitivity with increasing wave-length. This increase in effectiveness of the radiation parallels the opacity of the skin to the energy. When the white skin is painted with India ink the sensitivity to all wave-length bands is the same and equal to that of nonpenetrating infrared.

The negro subjects, although less sensitive to thermal stimuli, showed the same changes in the infrared regions as the whites, and it is only in the visible range that any significant effect due to pigment is seen. This difference in the visible spectrum is not sufficient, however, to make the negro appreciably more sensitive to sun light than the white man.

II. Dependence of sensation upon the surface area irradiated

The actual values of the stimuli given in Table II are not of themselves of any great interest since the area exposed to the radiation was chosen for convenience only. It has long been known that the sensation produced by a heat stimulus of fixed intensity increases with the area stimulated, but the matter has not received any systematic study due to a lack of a satisfactory technique. Radiation furnishes an ideal mode of stimulation for this investigation for the following reasons: (1) the strength of the stimulus can be readily changed and accurately measured; (2) the size of the area stimulated is limited only by the size of the experimental subject and can be easily measured; (3) the time of application of the stimulus can be measured and controlled; (4) the measurements of the stimuli are objective; (5) the skin temperature changes associated with any stimulus strength can be determined (see Section III).

TABLE III. *Minimum stimuli in g·cal./sec.·cm² for various sized areas on the body surface.*

Area		Minimum Stimuli			Skin Temperature Rise	
Location	Size	Subject I	Subject II	Average	Rate of Rise	3 Seconds
Forehead	0.95 cm²	0.0268 .0278	0.0294 .0204	0.0268	0.060°C/sec.	0.18°C
"	3.46	.0088 .0082 .0096	.0091 .0095	.0091	0.021	0.063
"	7.08	.0057 .0060	.0057 .0045 .0047	.0054	0.012	0.036
"	10.0	.0044 .0038	.0035 .0039	.0039	0.009	0.027
"	14.5	.0026	.0034 .0029 .0030	.0031	0.007	0.021
"	23.8	.0017	.0020 .0018	.0019	0.004	0.012
"	40.0	.0013	.0014 .0015	.0014	0.003	0.009
Entire Face	197	.00051 .00076	.00076 .00063	.00066	0.001+	0.004
Face and Chest	I 1680 II 1840	.0006 .0005	.0007 .0008 .0006	.0006	0.001+	0.004
Back to Waist	I 1680 II 1840	.0005 .0005	.0007 .0009	.00065	0.001+	0.004
Anterior Body Surface	I 5440 II 5590	.0004 .0005 .0004	.0006	.0005	0.001−	0.003−
Posterior Body	I 5440 II 5590	.0004 .0005	.0004 .0005	.00045	0.001−	0.003−

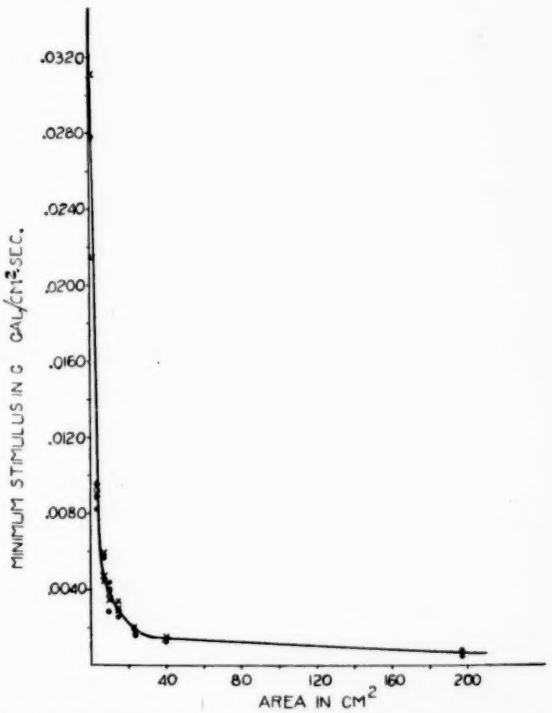


FIG. 3. Change in the energy required to evoke sensation with the area of the forehead exposed.

Of the three variables to be measured, namely, the magnitude of the sensation, stimulus strength, and surface area exposed, the first was eliminated as far as possible by keeping the sensation constant, i.e., at the minimum value. Thereby, the relationship between the stimulus strength and the area exposed, for a constant sensation, can be investigated. The experiments were done with nonpenetrating infrared radiation using the two white subjects of Section I. Far infrared was used in these tests because of its invisibility and the similarity of its stimulating qualities to other heat stimuli such as warm water or warm objects placed on the skin.

Because of its known uniform innervation the forehead was first studied. Starting with an area of 0.95 cm² the area exposed to the radiation was increased by steps using circular apertures of larger diameters up to 40 cm². The intensity of the radiation necessary to evoke a sensation for each size of the exposed area was measured in the manner described in Section I. In addition to the forehead, the whole face, the face and thorax, the anterior body surface, the back, and the entire posterior body surface were tested. The areas of

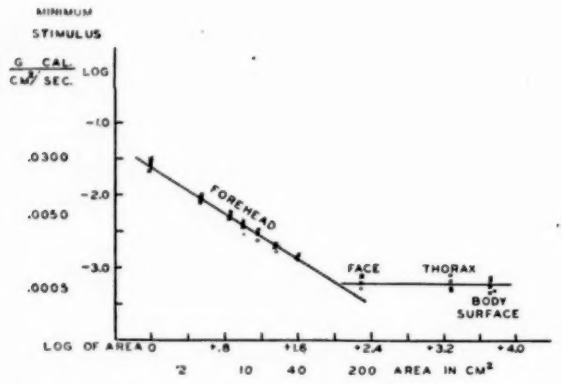


FIG. 4. Change in energy required to produce sensation with the size of body area exposed.

these larger surfaces were measured by means of a silhouette. The results of the tests are shown in Table III. Although the sensation results from the *thermal* changes in the skin, these changes are known (Section III) to be directly proportional to the intensity of the radiation. Therefore, in this discussion the radiation rate is used as a quantitative estimate of the stimulus strength.

The decrease in the amount of radiation required to produce a sensation as the exposed area is increased, confirms in a quantitative way the qualitative observations of others. The close agreement between the average values of the two subjects probably cannot be expected for every pair so tested, but the tests of the minimum stimuli made on a series of 12 subjects showed that a variation of more than 15 percent could usually be traced to lack of concentration or other disturbance. The values for the results on the forehead and face are shown in Fig. 3. The points fall on a smooth curve of the type of an equilateral hyperbola. The smoothness of the curve and the fact that the location of the test area on the forehead does not affect the stimulus value would lead one to conclude that the end organs' distribution over the forehead surface is effectively uniform. These facts also suggest that although the nerve ending itself is punctiform, it responds to stimuli over a considerable area. This is probably due to the circulation of blood in the skin area near the end organ. The close analogy between this arrangement and the usual physical thermopile with radiation receivers is obvious. It will be noticed that when an area as large as the face has been exposed, a sensitivity to heat has been reached which is not appreciably increased even by the exposure of the whole body. It is seen that relatively small portions of the body surface such as the face, arm, back, etc., are each as sensitive as the entire body surface. The extreme sensitivity of the skin to heat is evident from the fact that a rate of rise of skin temperature of less than a thousandth of a degree centigrade per second is perceived subjectively in three seconds. The whole of the data in Table III are represented graphically in Fig. 4.

The shape of the curve of Fig. 3 is of the usual exponential type and an analysis is easier in terms of the logarithms of the variables. Therefore the logarithms of the stimuli and of the areas are plotted in Fig. 4. The portion of the curve due to the forehead is a straight line. When the area is increased to include the face the line breaks sharply and becomes almost parallel to the area axis. Under no circumstances was a stimulus of less than 0.0004 g·cal./sec.·cm² observed regardless of the size of the area exposed.

Other portions of the body surface were examined in order to determine whether such a constant relationship as that found for the forehead could be found elsewhere. Tests were made on the arm and back which showed that unless counts of the end organs were made little could be learned.

From the results given in Table III it is evident that the sensation evoked by a stimulus is due to the added effects of the separate end organs, in much the same way that the voltage produced by a thermopile is increased with the number of junctions. In order to test further the combining power, or spatial summation, of the endings in various portions of the body the following experiments were performed. Two square apertures (25 cm² in area, 3 cm apart) were arranged so that the skin areas placed behind them could be stimulated singly and together. The procedure was to test each aperture alone and then together. Tests on the forehead showed that the right and left apertures gave identical values and together gave the value to be expected from Fig. 4. Tests were made of the backs of both hands, and of the forehead and the back of the right hand. The results are shown in Table IV.

The sensitivity of both hands together is seen

to be 50 percent greater than either hand alone, indicating that summation is taking place. The sensitivity of the combination of hand and forehead is no greater than the forehead alone, showing no summation.

III. The thermal changes in the skin accompanying heat sensation

From the foregoing it is evident that the temperature changes in the skin which accompany the minimal sensations are extremely small so that the usual procedure of following these changes with thermocouples either on the skin surface or buried in the tissue would not be likely to furnish reliable information. The network of blood vessels in the highly heat impervious tissue would invalidate the interpretation of thermal gradients from thermocouple measurements made at points deep in comparison with those at which the end organs are located. The presence of the thermocouple wires in the tissue also alters the normal heat flow to a great extent, especially at the point at which the thermocouple measures the temperature. For these reasons it was felt that it was not technically feasible to make direct measurements of the internal temperature changes in the layer of vascular tissue 0.3 mm thick which provoke temperature sensation.

The current theories as to the thermal changes responsible for heat sensation are four in number and are, briefly: The Hering[11] hypothesis which assumes sensation is dependent on the absolute change in the skin temperature (ΔT); the Weber hypothesis[12] supposes the rate of temperature change (dT/dt) to be the effector; the Ebbecke hypothesis[13] which assumes sensation to be due to a change in the internal thermal gradients of the tissues within 0.5 mm of the skin surface; the fourth hypothesis is a modification of Ebbecke's theory proposed by Bazett and McGlone[14] who postulate change in blood temperature rather than in the tissue.

TABLE IV. *Summation stimuli for various body areas.*

Backs of Both Hands

Subject I				Subject II			
Left	Right	Both	Summation	Left	Right	Both	Summation
0.0029 .0026	0.0039 .0025	0.0020 .0018 .0018 .0020	47%	0.0039	0.0041	0.0027	48%

Back of Right Hand and Forehead

Forehead	Hand	Both	Summation	Forehead	Hand	Both	Summation
0.0016	0.0035	0.0016	0%	0.0019	0.0039	0.0019	0%

[11] Hering, Sitzungsb. der Koeniges Akad. d. Wissensch., Wien **75**, 101 (1877).

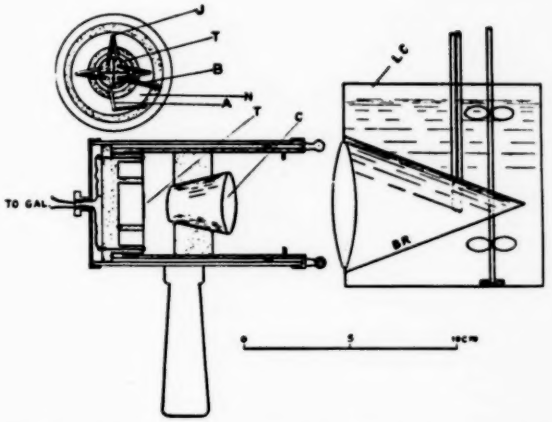


FIG. 5. Radiometer device used for measurement of skin temperature. Vertical section of radiometer (lower left); Leslie Cube (right), and detail of thermopile construction (upper right).

The present experiments were designed first with the idea of determining the magnitude of the skin temperature changes accompanying minimal sensations of heat, and second, by means of changing the internal thermal gradients with radiations of different penetrability, of throwing some light upon the questions of the thermal changes actually responsible for the sensation of warmth.

The apparatus for measuring the skin temperature is shown in Fig. 5. The method consists of comparing the radiation of the skin with that from an experimental black body.[15] By this means, the skin surface is not touched, the normal heat loss of the skin is not altered, nor is the skin compressed or indented with conse-

[12] E. H. Weber and R. Wagner, *Handwörterbuch der Physiologie* (Braunschweig, F. Vieweg und Sohn 1846), Vol. 3, p. 549.

[13] Ebbecke, Arch. f. d. ges. Physiologie **169**, 395 (1917).

[14] H. C. Bazett and B. McGlone, "Studies in Sensation III," Arch. Neurol and Psychiat. **28**, 91 (1932).

[15] J. D. Hardy, "Radiation of Heat from Human Body I," J. Clin. Invest. **13**, 593 (1934).

quent vascular changes as is the case when contact thermocouples and such devices are used. The surface of the white human skin emits radiation in the neighborhood of 9μ like a black body radiator,[16] so that no emissivity correction is necessary. The experimental black body is a Leslie Cube into one side of which has been built a conical cavity. The cone is painted black and serves adequately as a black body, the energy values determined therefrom agreeing within 1 percent with those obtained from a Radiation Standard of the National Bureau of Standards. During the experiments this black body was located near (Fig. 2) the skin surface to be measured and was kept at a constant temperature (33.5°C) by means of a thermostat. The instrument for detecting the radiation consisted of an eight junction bismuth alloy thermopile, the sensitive elements of which were arranged in a circular disk 1 cm in diameter. The disk was blackened with bismuth black and was supported by a thin film of nitrocellulose. In front of the thermopile was arranged a silver cone 10 cm long in order to focus the radiation from the skin surface. The sensitivity of the device was such that a difference in temperature between the Leslie Cube and the skin surface of 0.01°C would cause a deflection of 1 mm on the galvanometer scale. The instrument required 7 seconds to reach equilibrium after being removed from one radiant source to another. The radiometer was mounted on the vertical shaft of a pendulum so that it could be swung very rapidly from its normal position in front of the Leslie Cube to a position facing the skin. For obvious reasons the temperature changes of the skin could not be followed during the irradiation of the skin so the temperature was measured immediately before and after irradiation.

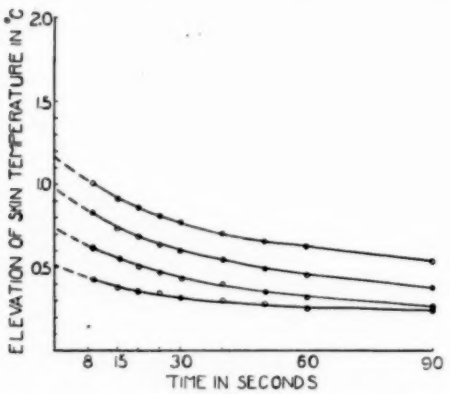


FIG. 6. Cooling curves of skin after irradiation showing extrapolation back to instant of shutting off radiation. Curve 1 = irradiation for 15 seconds; curve 2, for 30 seconds; curve 3, for 45 seconds; curve 4, for 60 seconds. Radiation = 0.0150 cal./sec.·cm² of nonpenetrating infrared.

[16] J. D. Hardy, "Radiation of Heat from Human Body III," J. Clin. Invest. **13**, 615 (1934).

It was soon discovered that the temperature changes accompanying minimal sensations of the large areas were too small to be measured directly, especially when it was considered that the change was always perceptible after 3 seconds of irradiation. For these reasons the temperature changes with larger intensities and long times of irradiation were measured and extrapolation resorted to for the smaller intensities and for the short times.

The arrangement for controlling and measuring the intensity of the source of radiation was the same as that employed for the measurement of the minimum stimuli. The forehead of the subject was used as a test area as its temperature is more constant than that of any other part of the body surface. The radiometer R_2 is the measuring device for skin temperature and the Leslie Cube is shown in its relation to the skin surface to be measured. The head of the subject was protected from the heat of the cube by a layer of insulation. The temperature of the cube was read by a sensitive thermometer readable directly to 0.005°C.

The subject was required to sit for about half an hour in the room before measurements were begun in order that his skin might come into equilibrium with the prevailing atmosphere. When it was found possible to obtain constant forehead temperature (within a few hundredths of a degree) the measurements were begun. The subject's forehead was then irradiated at a known rate for a fixed time. Immediately following the irradiation the radiometer was swung over to the skin surface and the cooling curve to the skin followed for 120 seconds. The intervals chosen for irradiation were 15, 30, 45 and 60 seconds. The type of curve obtained is shown in Fig. 6.

By extrapolation the cooling curves backward 8 seconds, the elevation of skin temperature due to the irradiation can be read. Such cooling

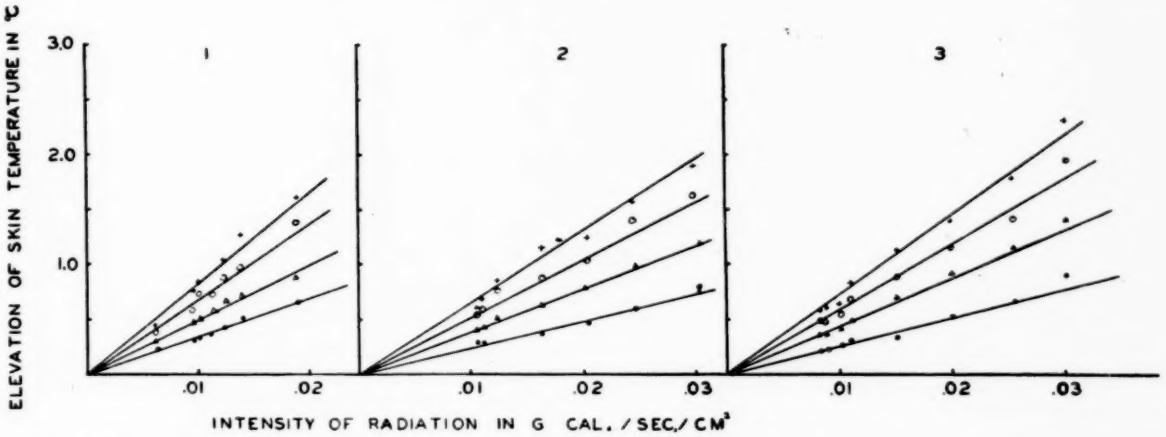


FIG. 7. Heating of the skin with the various radiations for constant times of irradiation. Dots = 15 seconds exposure; triangles, 30 seconds; circles, 45 seconds; crosses, 60 seconds. No. 1 = nonpenetrating infrared; No. 2 = penetrating infrared; No. 3 = visible radiation.

curves were plotted for various intensities of visible, near infrared and far infrared radiation. The temperature changes thus obtained were used to plot the heating curves of the skin for constant irradiation as shown in Fig. 7.

The direct proportionality between the energy absorbed by the skin and the rise in temperature is evidence of the fact that possible vasomotor and environmental factors did not enter in as serious errors of measurement. It is felt therefore that extrapolation from the experimental range of energies to smaller energies is valid.

The tests for sensitivity of the whole body were made only with nonpenetrating infrared radiation. Interpolation of the curves of Fig. 8 gives the temperature values shown in Table III.

The skin temperature elevations which occur when the skin absorbs the same number of calories of nonpenetrating infrared, penetrating infrared, and visible radiation is shown in Fig. 8. It is evident that the far infrared is the most efficient in raising the skin temperature, corresponding to the facts that it is most efficient in stimulating sensation and is totally absorbed within 0.1 mm of the skin surface. The thermal effects of the other radiations also correspond to their stimulating and penetrating qualities. In order to determine the skin temperature elevations by the radiations which produce the same sensation careful measurements were made of minimum stimuli, using an area of 7.08 cm². This area was chosen because the energy values are well within the experimental range of the temperature measurements. The results are shown in Table V. The average values are believed to be sufficiently consistent to be used for separating the effects due to the difference in penetration. The heating curves plotted from these energies are recorded in Fig. 9. It is evident that an equal sensation does not mean an equal rate of rise of temperature. Conversely, if the intensities of the three sources be adjusted so as to produce the same skin temperature rise, it will be found that the greatest sensation is produced by the far infrared, next by the near infrared, and next by the visible light. Thus, equal sensation does not mean an equal rise in skin temperature.

DISCUSSION

The dependence of heat sensation upon wavelength of the stimulating radiation is shown to be a matter of the optical properties of the skin alone, that is, of the reflectivity and the extinction coefficient of the skin for visible, penetrating infrared, nonpenetrating infrared radiation.[17] All differences in the stimulating effects are removed by covering the skin with black paint. Natural pigment, however, is seen to change the properties of the skin for visible radiation only, so that no idea of pigment function can be had by studying the effects of such artificial pigments as lamp black. The effect of natural skin pigment

[17] P. C. Foster, "Effect of Infrared on Tissue Temperature Gradient as Influenced by Pigment," Proc. Soc. Exper. Biol. and Med. **33**, 62 (1935).

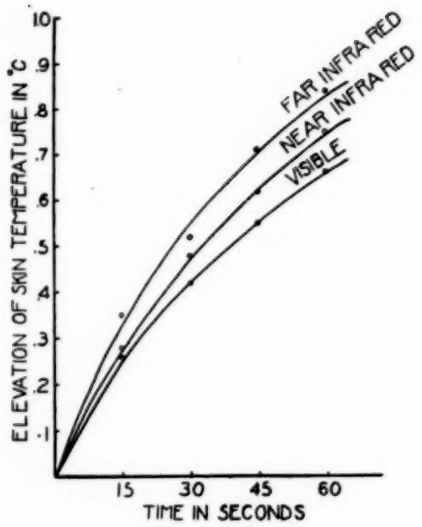


FIG. 8. Skin temperature change with absorption of equal energy of visible light, penetrating and nonpenetrating infrared.

upon the response of man to his thermal environment seems to be negligible so that negroes are not appreciably more sensitive to sunlight than whites and they respond in exactly the same way as do whites to infrared radiation. Pigment would seem therefore to have nothing to do with the regulation of heat exchanges between man and his environment, for not only does it not affect the sensitivity of the body to radiant heat but it does not aid in any way the elimination of heat from the body surface. Thus pigment affords no protection against the thermal effects of the natural sources of radiation and its protective action against the specific effect of ultraviolet radiation is doubtful, since the layers of the skin superficial to the pigment absorb practically all of the erythemal and antirachitic radiations.[18] Although pigment *formation* is undoubtedly stimulated by the ultraviolet radiation, *the effect* of the pigment seems to lie entirely within the visible range.

The straight line relationship for the logarithms of the values of stimulus and area for the forehead is so definite that it would seem to justify an attempt at relating these observations to sensation. That is, has a constant sensation definite meaning in terms of stimulus strength and

[18] W. W. Coblenz and R. Stair, "Transmissive Properties of Eye Protective Glasses and Other Substances," Technologic Papers of the National Bureau of Standards No. 369 (1928).

TABLE V. *Minimum stimuli of absorbed energy in g·cal./cm²/sec.*
Exposed Area = 7.08 cm².

Subject	Visible		Near Infrared		Far Infrared	
	I	II	I	II	I	II
Trial 1	0.0101	0.0109	0.0074	0.0086	0.0055	0.0060
2	.0112	.0114	.0074	.0087	.0052	.0052
3	.0115	.0111	.0088	.0081	.0064	.0052
4	.0112	.0112	.0087	.0078	.0045	.0052
5	.0117	.0116	.0087	.0078	.0054	.0052
Average	.0111	.0112	.0081	.0081	.0050	.0053

number of sensitive endings exposed? Although many of the physiological processes taking place between the stimulation of a sensory ending and the perception of sensation have been outlined by others, a complete picture of the sensory mechanism must necessarily include a study of the subjective impressions related to these processes. In line with modern thought, that sensation arises as the result of volleys of nerve impulses conducted to the brain by afferent nerve fibers from stimulated end organs, the following correlation is made between the present stimuli data of the forehead and the magnitude of sensation.

It is known that the *frequency* of the nervous discharge from an end organ increases with the strength of stimulation,[19] provided the stimulus be greater than the threshold of excitation of the

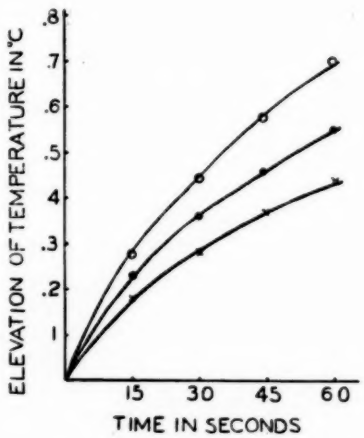


FIG. 9. Change in skin temperature by heating with radiation of visible light (circles), penetrating infrared (dots) and nonpenetrating infrared (crosses) which will produce the same sensation.

[19] E. D. Adrian and Y. Zotterman, "The Impulses Produced by Sensory Nerve Endings III." J. Physiol. **61**, 465 (1926).

ending, a phenomenon which is rather similar to the reaction of a photoelectric cell when exposed to light. The simplest assumption in regard to the relationship of sensation, stimulus strength, and frequency of nervous discharge is that the magnitude of the sensation is proportional to the impulse frequency reaching the brain, and the addition or association of the impulses into sensation. In the present experiment there were two conditions related to the sensation: (1) the sensation was constant in magnitude, and (2) the magnitude was that of the minimal sensation.

Thus, Minimum sensation = constant
= function (impulse frequency, association) (1)

is the relationship to be expected, on the basis of the above assumptions, from the present experiments. Of the factors in Eq. (1), the impulse frequency will depend upon: (1) the magnitude of the stimulus, (2) number of end organs stimulated, (3) the sensitivity of the individual endings. Factor (1) has been studied by Hartline,[20] Mathews[21] and others. Mathew's work on a cat's muscle shows that in the range of small stimuli the impulse frequency is roughly proportional to the logarithm of the stimulus strength. In accordance with Section II, the assumption is made that the end organ distribution over the forehead surface is uniform. That the sensitivity of all the endings is approximately the same cannot be definitely established by the present experiments but is evidenced by the fact that the minimum stimuli for such areas as the face, the arm and hand, the chest, the back, etc. are all the same. Thus, the total number of impulses arising from a stimulated area on the forehead would be proportional to $\log \varphi \times A$, φ being the intensity of the radiation and A the area stimulated. That this does not represent a constant sensation is seen from Table VI, where the tabulated values of $\varphi \times A$ are seen to increase with area size. The factor of central association or spatial summation must then be considered. By "spatial summation" is meant the addition in the brain of the separate effects of each end organ in the production of sensation.

The data of the present experiment can be accurately represented by the formula (see values of $A^{0.78} x \varphi$),

$$\log \varphi + a \log A = C \qquad (2)$$

where a and c are constants whose values are 0.78 and -1.502, respectively. The constant "c" can be associated directly with the magnitude of the sensory response, for as pointed out above, impulse frequency is known to be proportional to $\log \varphi$. The constant "a" we associate with summation of the separate end organs effects. Were the value of "a" unity, summation would be perfect and a constant stimulus magnitude would represent a constant sensory response. The value of "a" has two interpretations neither of which can be definitely established by these experiments. Either the number of end organs per unit area decreases systematically[22] with decrease in stimulus intensity or summation fails progressively with smaller intensities. The evidence in this experiment indicates that summation is the more important factor, and the minimum values of the summation are shown in Table VI.

Comparing Eq. (2) with the known laws of vision[23, 24] which are,

sensory response
$$= a \log \varphi + b \text{ (constant area)} \qquad (3)$$

sensory response
$$= c \log A + d \text{ (constant intensity)} \qquad (4)$$

shows that Eq. (3) (Weber-Fechner law) and 4 (Granit-Harper law) apply equally well to heat sensation. This identity affords gratifying evidence of the fundamental similarity of the peripheral and optical neurophysiological function.

TABLE VI.

Area (cm^2)	Stimulus (cal./sec.·cm^2)	$A \times \phi$	$A^{0.78} \times \phi$	Percent Summation
0.95	0.0268	0.025	0.026	
3.46	.0091	.032	.024	75
7.08	.0054	.038	.025	66
10.0	.0042	.042	.025	59
14.5	.0031	.045	.025	55
23.8	.0017	.045	.021	61
40.0	.0014	.056	.025	44

[20] H. K. Hartline and C. H. Graham, J. Cellular and Comparative Physiol. **5**, 277 (1932).
[21] B. H. C. Mathews, J. Physiol. **71**, 64 (1931).
[22] Selig Hecht, J. Gen. Physiol. **11**, 255 (1927).
[23] R. Granit and P. Harper, Am. J. Physiol. **95**, 211 (1930).
[24] G. T. Fechner, Abhandl. K. Sachs. Ges. Wissensch., math-phys. Kl. **4**, 31 (1858).

It would seem, therefore, that an interpretation of sensation can be made in terms of the stimulus and number of end organs. The relationship appears to be that the magnitude of the sensation depends upon the total frequency of impulses reaching the brain as modified by central association, and an equal sensation means an equal number of associated impulses. No value of association as high as 100 percent has been observed, and the magnitude of this variable depends probably upon habit. The function of summation in these experiments is analogous to the electrical resistance in a thermopile circuit. Perfect summation would mean a physiological thermopile of no resistance; the falling off of summation with increased numbers of end organs is therefore the physiological simile to the increased thermopile resistance with the number of junctions.

The difference between the threshold of sensation and the threshold of excitation for an end organ is clearly demonstrated by these experiments. The threshold of excitation for a single end organ is 0.0004 cal./sec.·cm², whereas that for sensation is more than 100 times as great. With relatively small areas a minimum stimulus in the neighborhood of 0.0004 cal./sec.·cm², is reached and when the area is increased beyond such a point no change in sensitivity is realized, that is, the minimum stimulus is constant. The sensation, however, increases as the exposed area is enlarged so that although the sensation is no longer minimal any decrease in the stimulus causes the sensation to disappear. Thus, the limit of sensitivity of the body is set not by degree of association of impulses but by the threshold of excitation of the end organ.

The temperature changes in the skin which produce sensation are seen from Section III to depend not only upon the heat absorbed at the skin surface but also upon the amount of heat absorbed in the various layers of skin below the surface. Thus, neither the actual change in surface temperature nor its rate of change can be assumed to be thermal changes responsible for heat sensation. Therefore, the Hering and Weber hypotheses may be ruled out. Body tissue itself is a very good insulator and this fact, together with the results given in Section III for small areas, which indicate that the end organs act as though supplied with a highly conducting receiving area, would rule out the possibility that the endings respond to changes in tissue temperature alone or that such changes are usually the important ones. Ebbecke's unmodified hypothesis can thus be said to be inadequate, although it is evident that his postulate in regard to a change in the internal thermal gradients being responsible for sensation is upheld by the present experiments. The hypothesis which comes nearest to explaining the present data is that the thermal receptors are really differential end organs, one sensitive spot located in a vascular bed superficial to the other, both elements being in close proximity to and surrounded by blood vessels. A differential change in the temperature of these blood vessel net works is assumed to be the necessary stimulus. A somewhat similar theory was proposed by Bazett and McGlone although their results seem to support a chemical rather than a thermal gradient as the stimulus.

The above interpretation is open to the criticism that no actual measurements were made of the temperature changes below the skin surface. As we have stated, we do not believe that significant measurements of these changes can be made directly. There is ample evidence, however, that the thermal gradients are changed differently by the types of radiation used in the above experiments and differences in the thermal sensitivity of the skin are found to correspond to these changes in thermal gradients. There is one phenomenon associated with temperature sensation which is particularly contradictory to any gradient hypothesis. Bazett found that when a thin fold of skin, containing end organs on only one side, was stimulated from the opposite side, a warm stimulus produced a warm sensation and a cold stimulus a cold sensation. On the basis of the gradient hypothesis paradoxical sensations would be expected such as are experienced when the skin is stimulated with a cool stimulus after having been subjected to a cold environment for some time. Since we have not repeated Bazett's experiment, our results must be considered as only confirmatory evidence for the gradient theory.

The magnitude of the temperature changes which produce minimal subjective sensations of

warmth are shown in Table III. A total rise in skin temperature of less than 0.003°C produced at a rate of 0.001°C per second is the sensitivity of the body as a whole and of many of its larger parts. Such sensitivity is evidence of the remarkable detector mechanism for the thermal regulator of the body temperature, and compares favorably with the most delicate of artificial thermoregulators. It is interesting to note the effect of habit upon the sensitivity of parts of the body. The parts of the skin surface usually most exposed to thermal changes are the face and hands and these parts although relatively small in area have a sensitivity almost equal to that of the whole body.

The radiation sensitivity of the skin (5×10^{-4} g·cal./sec.·cm^2) is not great when compared to that of a modern vacuum thermopile which can detect 10^{-12} g·cal./sec.·cm^2. It does compare favorably with the usual Coblenz type thermopile which is used in air. The thermopile used in the above experiments to measure the radiant energy, produced about 2.5 drum divisions at the minimum stimulus. A change in the temperature of the walls of a room of 3°C (that is a 30 percent change in the normal radiation rate), air temperature, velocity, humidity, etc., being constant, will evoke a sensation in 3 seconds. The energy exchange in this process is one-thirty millionth of the normal hourly radiation loss. That a sensation need not be necessary to this regulatory process is evident as end organ impulses can go to the brain without evoking a sensation. These impulses may have connection with the vasomotor pathways so that adjustment of internal thermal gradients of the skin can be made unconsciously, thereby decreasing end organ activity. In this way the small changes in environmental conditions can be automatically compensated for, and warning given of the larger changes which can be guarded against by the addition of clothing, etc.

Author Index to Volume 7

References with (L) are to Letters to the Editor.

Alcock, E. D. Effect of an electric field on the viscosity of liquids—126

Barnett, C. E. Effect of pigment particle size on some physical properties of rubber compounds—189
Barrett, C. S. and M. Gensamer. Stress analysis by x-ray diffraction—1
Baum, L. A. H. (see Traxler, R. N.)—9
Beams, J. S. (see Workman, E. J.)—375
Beckerley, J. G. (see Hansen, W. W.)—220
Beggs, J. Stiles. Synthesis of gear tooth curves—163
Bingham, Eugene. Book review—77
Bossart, P. N. Spectral emissivities, resistivity, and thermal expansion of tungsten-molybdenum alloys—50
Botset, H. G. (see Wyckoff, R. D.)—325
Braun, Milton L. Preliminary study of the after-effect or drift in rubber under constant load—421

Cady, Walter G. Piezoelectric resonator and the effect of electrode spacing upon frequency—237
Castleman, R. A., Jr. Hydrodynamical theory of piston ring lubrication—364
Cobine, J. D. Low pressure arc characteristics—137
Cookson, John W. and Harold Osterberg. Formula for the birefringence of vibrating media—166(L)
Cooper, Frank L. Atmospheric potential gradient anomalies —387
Cox, Martha (see Michels, Walter C.)—153

Davis, C. W. Some observations on the movement and demagnetization of ferromagnetic particles in alternating magnetic fields—236(L)
Dillon, J. H. Simplified extrusion plastometer for unvulcanized rubber—73
Dinnis, G. (see de Waele, A.)—426
Durand, Milo A. Coefficient of thermal expansion of magnesium oxide—297

Foote, Paul D. How physics is applied in the oil industry—91
Francis, Gladys M. Variations of the rigidity modulus with changes in applied stress—160
Frosch. C. J. VIII. Correlation of the distillation range with the penetration of creosote into paper strips—167

Gardner, Wm. Howlett. Plastic properties of shellac—306
Gemant, Andrew. Method of analyzing experimental results obtained from elasto-viscous bodies—311
Gensamer, M. (see Barrett, C. S.)—1
Goodier, J. N. Thermal stress in a strip due to variation of temperature along the length and through the thickness —156

Hansen, W. W. Directional characteristics of any antenna over a plane earth—460
——— **and J. G. Beckerley.** Radiation from an antenna over a plane earth of arbitrary characteristics—220
Hardy, James D. and Theodore W. Oppel. Thermal response of the skin to radiation—466
Haskell, N. A. Motion of a viscous fluid under a surface load. Part II—56
Hatfield, H. Stafford. Action of alternating and moving magnetic fields upon particles of magnetic substances—55(L)
Havenhill, R. S. Heat generation and anisotropy of rubber compounds in the flexometer—179
Hector, L. G. and H. L. Schultz. Dielectric constant of air at radiofrequencies—133
Hemingway, Allan and Glenn H. Witts. High frequency apparatus for producing a measured heat stimulus in studies of physiological temperature regulation—299
Hersey, Mayo D. Note on heat effects in capillary flow—403
Hicks, Victor. What x-rays can do for industry—79

Johnson, E. A. Design of tuned resistance-capacity coupled amplifiers—130

Keller, K. S. (see Pietenpol, W. B.)—43
Kent, R. H. Some special solutions for the motion of the powder gas—319
Kiehl, H. R. Electrical conductivity of glass. Part III. Current-voltage relationships with highly resistant layers—20
Koller, Lewis R. Cathode sputtering in arc discharges—225
——— Method for studying phenomena at the cathode in arc discharges—295
Kron, Gabriel. Quasi-holonomic dynamical systems—143

Labout, J. W. A. (see Saal, R. N. J.)—408

Macdonald, P. A. Thermionic amplification of direct currents—265
March, H. W. Bending of a centrally loaded rectangular strip of plywood—32
Meres, Milan W. (see Muskat, Morris)—346
Michels, Walter C. and Martha Cox. Thermal conductivity of tungsten—153
Mindlin, Raymond D. Force at a point in the interior of a semi-infinite solid—195
Mooney, M. Rheology of raw rubber—413
Morton, G. A. and E. G. Ramberg. Electron optics of an image tube—451
Murdock, C. C. and E. E. Zimmerman. Polarization impedance at low frequencies—211
Muskat, Morris. Seepage flux under dams of extended base width and under coffer dams resting on permeable strata of finite thickness. I—116
——— **and Milan W. Meres.** Flow of heterogeneous fluids through porous media—346
——— **and R. D. Wyckoff.** Theory of acid treatment of oil wells producing from limestone reservoirs—106

Oppel, Theodore W. (see Hardy, James D.)—466
Osterberg, Harold (see Cookson, John W.)—166

Pfeiffer, Hans H. Further tests of the elasticity of protoplasm—302
Pietenpol, W. B. Surface tension of molten glass—26
——— **and K. S. Keller.** Gassing of lead plate storage batteries during charge and discharge—43
Pugh, Emerson M. Some applications of physics to air conditioning—85

Ramberg, E. G. (see Morton, G. A.)—451
Rashevsky, N. Further studies on the mathematical physics of metabolizing systems with reference to living cells. II—260
Reiner, John M. Mechanism of division of metabolizing systems—368

Saal, R. N. J. and J. W. A. Labout. Relation between absolute viscosity and penetration of asphaltic bitumens —408
Saxl, Irving J. Application of chain weight loading to the determination of jell strength—62
Schultz, H. L. (see Hector, L. G.)—133
Schweyer, H. E. (see Traxler, R. N.)—67
Snoddy, L. B. (see Workman, E. J.)—375
Starr, Chauncey. Copper oxide rectifier—15
Sturtevant, Julian M. Heat transfer problem of interest in calorimetry—232

Tate, John T. Editorial—433
Traxler, R. N. and L. A. H. Baum. Permeability of compacted powders. Determination of average pore size—9
——— **and H. E. Schweyer.** Rheological properties of asphalts. III. A viscosity index—67

Ulrey, Dayton. New electronic tubes and new uses—97

de Waele, A. and G. Dinnis. Double mobility of some non-Newtonian fluids with particular reference to cellulose nitrate sols—426
Wells, D. A. Motion of a coil spring having variable tension and variable mass per unit length—395
Williams, Dudley. Application of infrared spectroscopy to rubber chemistry—399
Witts, Glenn H. (see Hemingway, Allan)—299
Wolff, Irving. Study of polarization capacity and electrode condition—203
Workman, E. J., J. W. Beams and L. B. Snoddy. Photographic study of lightning—375
Wyckoff, R. D. and H. G. Botset. Flow of gas-liquid mixtures through unconsolidated sands—325
——— (see Muskat, M.)—106

Yager, William A. Distribution of relaxation times in typical dielectrics—434
Young, Gaylord. Concentration fields and mechanical forces in metabolizing systems—380

Zimmerman, E. E. (see Murdock, C. C.)—211

Analytic Subject Index to Volume 7

References with (L) are to Letters to the Editor.

Alloys (see **Crystalline state**)
Amplifiers (see **Methods and instruments**)
Arcs (see **Discharge of electricity in gases**)

Batteries
Gassing of lead plate storage cells, W. B. Pietenpol, K. S. Keller—43
Biophysics
Elasticity of protoplasm, H. H. Pfeiffer—302
High frequency apparatus for heat stimulus, A. Hemingway, G. H. Witts—299
Metabolizing systems, concentration fields and mechanical forces, G. Young—380
Metabolizing systems, division of, J. M. Reiner—368
Physics of metabolizing systems, N. Rashevsky—260
Response of skin to radiation, J. D. Hardy, T. W. Oppel—466
Book reviews
First report on viscosity and plasticity, J. M. Bürgers, et al—77

Crystalline state
Spectral emissivities, resistivity, and thermal expansion of W-Mo alloys, P. N. Bossart—50

Dielectric constants
Of air at radio frequencies, L. G. Hector, H. L. Schultz—133
Dielectrics and dielectric properties
Relaxation times, W. A. Yager—434
Discharge of electricity in gases
Arc characteristics at low pressures, J. D. Cobine—137
Photographs of lightning, E. J. Workman, J. W. Beams, L. B. Snoddy—375
Sputtering in arc discharges, L. R. Koller—225
Temperature of cathode, L. R. Koller—295
Dynamics
Quasi-holonomic dynamical systems, G. Kron—143

Elasticity
Anisotropy of rubber compounds in flexometer, R. S. Havenhill—179
Bending of plywood, H. W. March—32
Determination of jell strength, I. J. Saxl—62
Drift in rubber under constant load, M. L. Braun—421
Elasto-viscous bodies, behavior of, A. Gemant—311
Force in interior of a solid, R. D. Mindlin—195
Motion of coil spring; variable tension and density, D. A. Wells—395
Piezoelectric resonator, W. G. Cady—237
Pigment particle size and properties of rubber, C. E. Barnett—189
Plastic properties of shellac, W. H. Gardner—306
Plastometer for unvulcanized rubber, J. H. Dillon—73
Properties of asphaltic bitumens, R. N. J. Saal, J. W. A. Labout—408
Of protoplasm, H. H. Pfeiffer—302
Rheology of raw rubber, M. Mooney—413
Rigidity modulus and changes in applied stress, G. M. Francis—160
Stress analysis by x-rays, C. S. Barrett, M. Gensamer—1
Stress in a strip due to variation of temperature, J. N. Goodier—156
Electrical circuits
Amplification of direct currents, P. A. Macdonald—265
High frequency apparatus for heat stimulus, A. Hemingway, G. H. Witts—299
Tuned resistance-capacity coupled amplifiers, E. A. Johnson—130
Electrical conductivity and resistance
Of glass, H. R. Kiehl—20
Of W-Mo alloys, P. N. Bossart—50
Electrolytic cells
Polarization capacity and electrode condition, I. Wolff—203
Polarization impedance, C. C. Murdock, E. E. Zimmerman—211
Electromagnetic theory
Directional characteristics of antenna, W. W. Hansen—460
Radiation from antenna, W. W. Hansen, J. G. Beckerley—220
Electron optics
Aberrations and focusing properties of image tubes, G. A. Morton, E. G. Ramberg—451
Emissivities
Of W-Mo alloys, P. N. Bossart—50

Fluidity (see **Viscosity**)

Gases (see **Kinetic theory**)
Geophysics
Motion under load on earth's surface, N. A. Haskell—56

Hydrodynamics
Flow of fluids through porous media, M. Muskat, M. W. Meres—346
Flow of gas-liquid mixtures through sands, R. D. Wyckoff, H. G. Botset—325
Heat effects in capillary flow, M. D. Hersey—403
Piston ring lubrication, theory, R. A. Castleman, Jr.—364
Seepage flux under dams, M. Muskat—116

Instruments (see **Methods and instruments**)

Kinetic theory of gases
Motion of powder gas, R. H. Kent—319

Liquids (see **Hydrodynamics**)
Double mobility of non-Newtonian fluids, A. de Waele, G. Dinnis—426

Liquids (see **Hydrodynamics**) (continued)
Seepage flux under dams, M. Muskat—116
Viscous fluid under a surface load, N. A. Haskell—56

Magnetic properties
Alternating fields and particles, C. W. Davis—236(L); H. S. Hatfield—55(L)
Mechanics (see **Dynamics**)
Meteorology
Photographs of lightning, E. J. Workman, J. W. Beams, L. B. Snoddy—375
Potential gradient anomalies, F. L. Cooper—387
Methods and instruments
Air-conditioning, applications of physics, E. M. Pugh—85
Amplification of direct currents, P. A. Macdonald—265
Analysis of vibrations of transparent media, J. W. Cookson, H. Osterberg—166(L)
Cathode sputtering in arc discharges, L. R. Koller—225
Copper-oxide rectifier, C. Starr—15
Determination of jell strength, I. J. Saxl—62
Electronic tubes and new uses, D. Ulrey—97
Extrusion plastometer, J. H. Dillon—73
High frequency apparatus for heat stimulus, A. Hemingway, G. H. Witts—299
Oil industry, applications of physics, P. D. Foote—91
Phenomena at cathode, method for studying, L. R. Koller—295
Rheology of raw rubber, M. Mooney—413
Stress analysis by x-rays, C. S. Barrett, M. Gensamer—1
Synthesis of gear tooth curves, J. S. Beggs—163
Tuned resistance-capacity coupled amplifiers, E. A. Johnson—130
X-rays in industry, V. Hicks—79

Oil industry
Acid treatment of wells, M. Muskat, R. D. Wyckoff—106
Applications of physics, P. D. Foote—91
Optical theory
Birefringence of vibrating media, J. W. Cookson, H. Osterberg—166(L)

Permeability
Of compacted powders, R. N. Traxler, L. A. H. Baum—9
Piezoelectric effect
Resonator and effect of electrode spacing, W. G. Cady—237
Plasticity (see **Elasticity**)

Radio
Directional characteristics of antenna, W. W. Hansen—460
Radiation from antenna, W. W. Hansen, J. G. Beckerley—220
Piezoelectric resonator, W. G. Cady—237
Rectifiers
Copper-oxide; characteristics, C. Starr—15
Resistance, electrical (see **Electrical conductivity and resistance**)

Spectra, absorption
Infrared spectroscopy and rubber chemistry, D. Williams—399
Spectroscopy, technique
Infrared spectroscopy and rubber chemistry, D. Williams—399
Surface tension
Of molten glass, W. B. Pietenpol—26
Penetration of creosote into paper strips, C. J. Frosch—167

Thermal conductivity
Air-conditioning problems, E. M. Pugh—85
Heat effects in capillary flow, M. D. Hersey—403
Heat transfer in calorimetry, J. M. Sturtevant—232
Of tungsten, W. C. Michels, M. Cox—153
Thermal expansion
Of magnesium oxide, M. A. Durand—297
Stress in a strip due to variation of temperature, J. N. Goodier—156
Of W-Mo alloys, P. N. Bossart—50

Vacuum tubes
New uses, D. Ulrey—97
Viscosity
Double mobility of non-Newtonian fluids, A. de Weale, G. Dinnis—426
Drift in rubber under constant load, M. L. Braun—421
Effect of an electric field, E. D. Alcock—126
Elasto-viscous bodies, behavior of, A. Gemant—311
Penetration of asphaltic bitumens and viscosity, R. N. J. Saal, J. W. A. Labout—408
Plastic properties of shellac, W. H. Gardner—306
Rheology of raw rubber, M. Mooney—413
Temperature viscosity relations for asphalts, R. N. Traxler, H. E. Schweyer—67

X-rays, diffraction, scattering, reflection, refraction and polarization
Stress analysis by x-rays, C. S. Barrett, M. Gensamer—1
X-rays in industry, V. Hicks—79